

63548

**Delaval, Jan**

---

**From:** Gambel, Phillip  
**Sent:** Tuesday, April 02, 2002 9:15 AM  
**To:** Delaval, Jan  
**Subject:** exam answer 09 / 020393

jan

please perform a sequence and a sequence interference search for

ussn 09 / 020393

open and closed

amino acids 42-58 of SEQ ID NO: 3

thanx

phillip gambel  
art unit 1644  
308-3997

Room 8 B03

1644 mailbox

please let me know when it is ready. thanx.

Jan Delaval  
Reference Librarian  
Biotechnology & Chemical Library  
CM1 1E07 - 703-308-4498  
jan.delaval@uspto.gov

**THIS PAGE BLANK (USPTO)**

**SEARCH REQUEST FORM**

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
 Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
 Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

**If more than one search is submitted, please prioritize searches in order of need.**

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

Jan Delaval  
 Reference Librarian  
 Biotechnology & Chemical Library  
 CM1 1E07 - 703-308-4498  
 jan.delaval@uspto.gov

\*\*\*\*\*  
**STAFF USE ONLY**

Searcher: Jan  
 Searcher Phone #: 4498  
 Searcher Location: \_\_\_\_\_  
 Date Searcher Picked Up: 7/2/02  
 Date Completed: 7/2/02  
 Searcher Prep & Review Time: \_\_\_\_\_  
 Clerical Prep Time: 15  
 Online Time: 20

**Type of Search**

NA Sequence (#) \_\_\_\_\_  
 AA Sequence (#) 4  
 Structure (#) \_\_\_\_\_  
 Bibliographic \_\_\_\_\_  
 Litigation \_\_\_\_\_  
 Fulltext \_\_\_\_\_  
 Patent Family \_\_\_\_\_  
 Other \_\_\_\_\_

**Vendors and cost where applicable**

STN \_\_\_\_\_  
 Dialog \_\_\_\_\_  
 Questel/Orbit \_\_\_\_\_  
 Dr.Link \_\_\_\_\_  
 Lexis/Nexis \_\_\_\_\_  
 Sequence Systems ✓  
 WWW/Internet \_\_\_\_\_  
 Other (specify) \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**



63548  
**Delaval, Jan**

---

**From:** Gambel, Phillip  
**Sent:** Tuesday, April 02, 2002 10:07 AM  
**To:** Delaval, Jan  
**Subject:** second sequence to search for 09 /020393

jan

please perform a sequence and a sequence interference search

both open and closed

for

09 / 020393

SEQ ID NO: 18

thanx

phillip gambel  
art unit 1644  
308-3997

1644 mailbox 9E12

**Jan Delaval**  
**Reference Librarian**  
Biotechnology & Chemical Library  
CM1 1E07 – 703-308-4498  
jan.delaval@uspto.gov

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:19:51 : Search time 12.82 seconds  
(without alignments)  
101.011 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58

Perfect score: 96  
Sequence: 1 FEHCNFNDVYTRRENE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database : PIR\_68:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	128	1 RMH059	surface glycoprote
2	63	65.6	126	2 I36914	CD59 protein - bab
3	63	65.6	128	2 I36894	CD59 protein - gre
4	46.5	48.4	1302	2 T20767	hypothetical prote
5	46.5	48.4	1767	2 T20766	hypothetical prote
6	46	47.9	284	2 E83557	probable transcrip
7	46	47.9	777	2 E83748	hypothetical prote
8	45	46.9	1307	2 T21283	hypothetical prote
9	44	45.8	205	2 H71639	NADH dehydrogenase
10	44	45.8	807	2 D69102	collagenase - Meth
11	43	44.8	544	2 S41626	spike protein chat
12	42	43.8	365	2 T20652	hypothetical prote
13	42	43.8	419	2 C83681	ABC transporter (s
14	42	43.8	464	2 T16889	hypothetical prote
15	41	42.7	117	2 C69969	hypothetical prote
16	41	42.7	179	2 S23358	H+-transporting AT
17	41	42.7	223	1 VC8VCA	coat protein - tob
18	41	42.7	265	1 S64938	hypothetical prote
19	41	42.7	410	2 T51212	related to integra
20	41	42.7	720	2 T38647	hypothetical prote
21	41	42.7	742	2 T23226	hypothetical prote
22	41	42.7	805	2 T03896	hypothetical prote
23	41	42.7	906	2 T48898	disease resistance
24	41	42.7	908	2 T48898	disease resistance
25	41	42.7	1162	2 S07421	E2 glycoprotein pr
26	41	42.7	1162	2 S14939	E2 glycoprotein pr
27	41	42.7	1162	2 S14940	E2 glycoprotein pr
28	41	42.7	1528	2 S13743	DNA strand transfe
29	41	42.7	1646	1 WMTMS2	186k protein - cuc

30	41	42.7	3433	1 S28381	utrophin - human
31	40	41.7	78	2 H84010	hypothetical prote
32	40	41.7	165	2 F71691	hypothetical prote
33	40	41.7	218	2 F69972	probable membrane
34	40	41.7	244	2 T00825	probable heat choc
35	40	41.7	497	2 T06727	hypothetical prote
36	40	41.7	607	2 T43322	meiosis specific p
37	40	41.7	742	1 A43344	synaptic vesicle p
38	40	41.7	742	2 S27263	synaptic vesicle p
39	39.5	41.1	156	2 T31664	DAP-kinase homolo
40	39.5	41.1	169	2 B64374	hypothetical prote
41	39.5	41.1	632	2 T31667	guanylate cyclase
42	39.5	41.1	632	2 T31666	guanylate cyclase
43	39.5	41.1	760	2 S75517	natruietic peptid
44	39.5	41.1	854	1 VCLJST	env polyprotein pr
45	39	40.6	119	2 T32569	hypothetical prote
46	39	40.6	258	2 T13591	tail fiber adhesin
47	39	40.6	289	2 S73935	hypothetical prote
48	39	40.6	287	2 I38517	synthaxin - human
49	39	40.6	297	2 S52726	synthaxin-4 - human
50	39	40.6	332	2 C96693	hypothetical prote

## ALIGNMENTS

RESULT 1  
RMH059  
surface glycoprotein CD59 precursor [validated] - human  
N/Alternate names: 1F5 antigen protein; 20K homologous restriction factor (HRF20); CD  
plex Inhibition Factor (MACIF); membrane inhibitor of reactive lysis (MIRL); protecti  
C/Species: Homo sapiens (man)  
C/Date: 30-Sep-1990 #sequence, revision 30-Sep-1990 #text, change 08-Dec-2000  
C/Accession: A46252; JI0109; A33405; JI0134; A34587; S05504; S09201; A60828; PL0041;  
R/Betrunka, J.G.; Fleenor, D.E.; Sykes, K.; Kaufman, R.E.; Rosse, W.F.  
Proc. Natl. Acad. Sci. U.S.A. 89, 7876-7879, 1992  
A/Title: Structure of the CD59-encoding gene: Further evidence of a relationship to m  
A/Reference number: A46252; MIMD:92390353  
A/Accession: A46252  
A/Molecule type: DNA  
A/Residues: 1-128 <PEP>  
A/Cross-references: GB:M82840; NID:9180149; PIDN:AAA68793.1; PID:9180151  
A/Note: sequence extracted from NCBI backbone (NCBIN:112714, NCBIN:112718, NCBIN:1127  
R/Davies, A.; Simmons, D.L.; Hale, G.; Harrison, R.A.; Tighe, H.; Lachmann, F.J.; Wal  
J. Exp. Med. 170, 637-654, 1989  
A/Title: CD59, an LY-6-like protein expressed in human lymphoid cells, regulates the  
A/Reference number: JI0109; MIMD:89361238  
A/Accession: JI0109  
A/Molecule type: mRNA  
A/Residues: 1-128 <DAV>  
A/Cross-references: EMBL:X16447; NID:929805; PIDN:CAA34467.1; PID:929806  
R/Okada, H.; Nagami, Y.; Takahashi, K.; Okada, N.; Hideshima, T.; Takizawa, H.; Kondo  
Biochem. Biophys. Res. Commun. 167, 1553-1559, 1989  
A/Title: 20 kDa homologous restriction factor of complement resembles T cell activati  
A/Reference number: A33405; MIMD:89350963  
A/Accession: A33405  
A/Molecule type: mRNA  
A/Residues: 1-128 <OKA>  
A/Cross-references: GB:M27909; NID:9623406; PIDN:AAA60543.1; PID:9623407  
R/Sugita, Y.; Tohe, T.; Oda, E.; Tomita, M.; Yasukawa, K.; Yamaji, N.; Takemoto, T.;  
J. Biochem. 106, 555-557, 1989  
A/Title: Molecular cloning and characterization of MACIF, an inhibitor of membrane ch  
A/Reference number: JI0134; MIMD:90110046  
A/Accession: JI0134  
A/Molecule type: mRNA  
A/Residues: 1-128 <SUG>  
A/Note: parts of this sequence, including the amino end of the mature protein, were c  
A/Note: sites for glycosylation and the absence of glycosylation were confirmed  
R/Sawada, B.I.; Ohashi, K.; Anaguchi, H.; Okazaki, H.; Hattori, M.; Minato, N.; Naruto,  
DNA Cell Biol. 9, 213-220, 1990  
A/Title: Isolation and expression of the full-length cDNA encoding CD59 antigen of hu  
A/Reference number: A34587; MIMD:90253615  
A/Accession: A34587

A: Molecule type: mRNA  
 A: Residues: 1-128 <SAM>  
 A: Cross-references: GB:M34671; NID:g180152; PIDN:AAA51952.1; PID:g180153  
 R: Savade, R.; Ohashi, K.; Okano, K.; Hattori, M.; Minato, N.; Natuto, M.  
 A: Title: Complementary DNA sequence and deduced peptide sequence for CD59/MEW-43 antigen  
 A: Reference number: S05504; MUID:89386002  
 A: Accession: S05504  
 A: Molecule type: mRNA  
 A: Residues: 27-128 <SA2>  
 A: Cross-references: EMBL:X15861; NID:g29803; PIDN:CAA33870.1; PID:g1340180  
 R: Philbrick, W.M.; Palfree, R.G.E.; Maher, S.E.; Bridgett, M.M.; Sirlin, S.; Bothwell, A.  
 A: Title: The CD59 antigen is a structural homologue of murine Ly-6 antigens but lacks its  
 A: Reference number: S09201; MUID:90168959  
 A: Accession: S09201  
 A: Molecule type: mRNA  
 A: Residues: 1-128 <PHD>  
 A: Cross-references: EMBL:X17198; NID:g29814; PIDN:CAA35059.1; PID:g29815  
 R: Cabral, A.R.; Cole, L.A.; Walz, D.A.; Castor, C.W.  
 A: Title: Connective tissue activation, XXXII. Structural and biologic characteristics of  
 A: Reference number: A60828; MUID:88134429  
 A: Accession: A60828  
 A: Molecule type: protein  
 A: Residues: 26-27, 'Y', 29-30, 'D', 32-37, 'X', 39-42, 'XX', 45-50, 'X', 52-62; 'VXRLLD' <CAB>  
 A: Experimental source: normal urine  
 A: Note: the six unknown or mismatched residues in the amino-terminal fragment correspond  
 R: Stefanova, I.; Hilgert, I.; Kristofova, H.; Brown, R.; Low, M.G.; Horejsl, V.  
 A: Title: Characterization of a broadly expressed human leucocyte surface antigen MEW-43  
 A: Reference number: PL0041; MUID:89143489  
 A: Accession: PL0041  
 A: Molecule type: protein  
 A: Residues: 26-42 <SRP>  
 R: Harada, R.; Okada, N.; Fujita, T.; Okada, H.  
 A: Title: Purification of IFS antigen that prevents complement attack on homologous cell  
 A: Reference number: A60774; MUID:90171576  
 A: Accession: A60774  
 A: Molecule type: protein  
 A: Residues: 26-42, 'XX', 45-50, 'X', 52, 'X', 54-57, 'X', 59-63 <HAR>  
 R: Nishimura, H.; Stewart, B.H.; Rollins, S.A.; Zhao, J.; Bothwell, A.L.M.; Sims, P.J.  
 A: Title: Contribution of the N-linked carbohydrate of erythrocyte antigen CD59 to its CD  
 A: Reference number: A38089; MUID:92235065  
 A: Accession: A38089  
 A: Molecule type: protein  
 A: Residues: 40-42, 'X', 44-49 <NIN>  
 R: Sugita, Y.; Nakano, Y.; Oda, E.; Noda, K.; Tobe, T.; Miura, N.H.; Tomita, M.  
 A: Title: Determination of carboxyl-terminal residue and disulfide bonds of MAC1F(CD59),  
 A: Reference number: PX0068; MUID:94103166  
 A: Accession: PX0068  
 A: Molecule type: protein  
 A: Residues: 26-29, 30-39, 40-42, 'X', 44-48, 49-52, 56-65, 65-72, 88-90, 92-96 <SU2>  
 R: Tone, M.; Walsh, L.A.; Waldmann, H.  
 A: Title: Gene structure of human CD59 and demonstration that discrete mRNAs are generated  
 A: Reference number: I37223; MUID:93021133  
 A: Accession: I37223  
 A: Status: translated from GB/EMBL/DBJ  
 A: Molecule type: DNA  
 A: Residues: 1-128 <RES>  
 A: Cross-references: EMBL:Z14113; NID:g29810; PIDN:CAA78486.1; PID:g825637  
 C: Comment: This cell surface antigenic glycoprotein inhibits homologous complement lysis  
 C: Comment: of the binding of C9 and C8 to C5b  
 C: Comment: This glycosylphosphatidylinositol-anchored protein is deficient in cases of F  
 A: Genes: GDB:CD59  
 A: Cross-references: GDB:119769; OMIM:107271  
 A: Map position: 11p13-11p13

A: Introns: 23/1; 57/1  
 A: Note: The first intron occurs before the initiator codon  
 C: Superfamily: Ly-6 antigen; Ly-6 homology  
 C: Keywords: blocked carboxyl end; glycoprotein; lipoprotein; membrane protein; phosph  
 F: 26-102/Product: signal sequence status predicted: lipoprotein; membrane protein; phosph  
 F: 26-102/Product: surface glycoprotein CD59 #status experimental <MAT>  
 F: 103-128/Domain: Ly-6 homology <LY6>  
 F: 28-51, 31-38, 44-64/Disulfide bonds: #status experimental  
 F: 43/Binding site: carbohydrate (Asn) (covalent) #status absent  
 F: 70-88, 89-94/Disulfide bonds: (or 70-89, 88-94) #status experimental  
 F: 102/Modified site: GPI anchor ethanolamine amidated carboxyl end (Asn) (in mature  
 Query Match  
 Best Local Similarity 100.0%; Score 96; DB 1; Length 128;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2  
 I36914  
 CD59 protein - baboon  
 C: Species: Papio sp. (baboon)  
 C: Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999  
 C: Accession: I36914  
 R: Fodor, W.L.; Rollins, S.A.; Bianco-Caron, S.; Burton, W.V.; Guilmette, E.R.; Rother  
 Immunogenetics 41, 51, 1995  
 A: Title: Primate terminal complement inhibitor homologues of human CD59.  
 A: Reference number: I36894; MUID:95104908  
 A: Accession: I36914  
 A: Status: preliminary; translated from GB/EMBL/DBJ  
 A: Molecule type: DNA  
 A: Residues: 1-126 <RES>  
 A: Cross-references: GB:I22862; NID:g514327; PIDN:AAA74127.1; PID:g514328  
 C: Superfamily: Ly-6 antigen; Ly-6 homology  
 F: 26-100/Domain: Ly-6 homology <LY6>

Query Match  
 Best Local Similarity 65.6%; Score 63; DB 2; Length 126;  
 Matches 10; Conservative 5; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 FEHCNFDVYTRLRENE 17  
 Db 67 FANCFNDISTLKESE 83

RESULT 3  
 I36894  
 CD59 protein - green monkey  
 C: Species: Cercopithecus aethiops (green monkey, grivet)  
 C: Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999  
 C: Accession: I36894  
 R: Fodor, W.L.; Rollins, S.A.; Bianco-Caron, S.; Burton, W.V.; Guilmette, E.R.; Rother  
 Immunogenetics 41, 51, 1995  
 A: Title: Primate terminal complement inhibitor homologues of human CD59.  
 A: Reference number: I36894; MUID:95104908  
 A: Accession: I36894  
 A: Status: preliminary; translated from GB/EMBL/DBJ  
 A: Molecule type: DNA  
 A: Residues: 1-128 <RES>  
 A: Cross-references: GB:I22863; NID:g514314; PIDN:AAA74126.1; PID:g514315  
 C: Superfamily: Ly-6 antigen; Ly-6 homology  
 F: 26-102/Domain: Ly-6 homology <LY6>

Query Match  
 Best Local Similarity 65.6%; Score 63; DB 2; Length 128;  
 Matches 10; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

	Matches	10;	Conservative	5;	Mismatches	2;	Indels	0;	Gaps	0;
QY	1	FEHCNFNDVTRLRENE	17							
		: :    :	: :							
Db	67	FANCFNFDISTLIKESF	83							

RESULT 4  
T20767  
Hypothetical protein PflC1.5b - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T20767  
R:Palmer, S.  
submitted to the EMBL Data Library, September 1995  
A:Reference number: Z19321

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1302 <MIL>  
A:Cross-References: EMBL:554270; PIDN:CAA91031.1; GSPDB:GN00028; CESP:FL1CL.5b  
A:Experimental source: clone FL1CL  
C:Genetics:  
A:Gene: CESP:FL1CL.5b  
A:Map position: X  
A:Introns: 21/1, 97/3, 134/3, 189/3, 245/3, 284/3, 474/2, 553/3, 701/2, 734/3, 807/1, 88

Query Match	Score	DB	Length
Best Local Similarity	48.4%	46.5	1302
Matches	12	Conservative	1; Mismatches 4; Indels 5; Gaps 2
QY	1	FEH---- <td>ENE 17</td>	ENE 17
DB	601	FEHELNTANVEDTTRIKRNE	622

RESULT 5  
T20766  
hypothetical protein F11C1.5a - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T20766  
R:Palmer, S.  
submitted to the EMBL Data Library, September 1995  
A:Reference number: Z19321  
A:Accession: T20766  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1767 <MIL>  
A:Cross-references: EMBL:Z54270; PIDN:CAA91030.1; GSPDB:GN00028; CESP:F11C1.5a  
A:Experimental source: clone F11C1  
C:Genetics:  
A:Gene: CESP:F11C1.5a  
A:Map position: X  
A:Introns: 21/1: 97/3; 134/3; 189/3; 245/3; 284/3; 474/2; 553/3; 701/2; 734/3; 807/1; 889/3; 1653/3; 1700/2

Query Match	48.4%	Score 46.5	DB 2	Length 1767
Best Local Similarity	54.5%	Pred. No. 31		
Matches 12	Conservative 1	Mismatches 4	Indels 5	Gaps 2
QY	1 FEH-----CN-FNDVTRLRNE	17		
Db	601 FEHELNTANVFDDTIRKENE	622		

```

RESULT      6
EB3557
probable transcription regulator PA0708 [imported] - Pseudomonas aeruginosa (strain PA01)
c:Species: Pseudomonas aeruginosa
c:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

```

C:Accession: E83557  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; ...  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardvig, K.; L...  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa...  
A:Reference number: AB2950; MUID:20437337  
A:Accession: E83557  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-784 <STO>  
A:Cross-references: GB:AE004506; GB:AE004091; NID:99946584; PIDN:AG04097.1; GSPDB:GN...  
A:Experimental source: strain PA01  
A:Genetics:  
A:Gene: PA0708

Query Match	47.9%	Score 46	DB 2	Length 284
Best Local Similarity	60.0%	Pred. No. 5.5		
Matches	9	Conservative	3	Mismatches 3; Indels 0; Gaps 0.
Qy	3	HCNENDVTRLPENE	17	
Db	26	HCVPSNITTRLE	40	

```

RESULT      7
EB3748
Hypotheoretical protein BH0789 [imported] - Bacillus halodurans (strain C-125)
C|Species: Bacillus halodurans
C|Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
C|Accession: EB3748
R|Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A|Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A|Reference number: A83650; MUID:20263314
A|Accession: EB3748
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1*777 <SMO>
A|Cross-references: GB:AP001509; GB:BA000004; NID:g10173176; PIDN:BA04508.1; GSPDB:G
A|Experimental source: strain C-125
C|Genetics:
A|Gene: BH0789

```

Query Match	47.9%	Score 46;	DB 2;	Length 777;
Best Local Similarity	46.7%	Pred. No. 16;		
Matches	7;	Conservative	5;	Mismatches 3; Indels 0; Gaps 0;
OY	1	FEHCNFENDYTRFLRE	15	
Db	136	FKHSNFHDYLTQIKE	150	

RESULT 8  
T21283  
hypothetical protein F23A7.5 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T21283  
R:McMurray, A.  
submitted to the EMBL Data Library, October 1996  
A:Reference number: Z19401  
A:Accession: T21283  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1307 <MIL>  
A:Cross-references: EMBL:Z81067; PIDN:CAB02976.1; GSPDB:GN00028; CESP:F23A7.5  
A:Experimental source: clone F23A7  
C:Genetics:  
A:Gene: CESP:F23A7.5  
A:Map position: X

A;Introns: 154/1; 230/2; 271/1; 307/1; 466/2; 551/2; 627/3; 663/1; 699/3; 746/3; 772/3;

## Query Match

Best Local Similarity 46.9%; Score 45; DB 2; Length 1307;  
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 4 CBNNDVTTLREN 16  
654 CKINDDTTLREN 666

## RESULT 9

H71639

NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain J RP790 - Rickettsia prowazekii

C;Species: Rickettsia prowazekii

C;Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000

R;Accession: H71639

R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alismark, U.

A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A;Reference number: A71630; MUID:99039499

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Residues: 1-205 <AND>

A;Cross-references: GB:AJ235273; GB:AJ235269; MID:93861237; PIDN:CAAL5216.1; PID:9386131

A;Experimental source: strain Madrid E

A;Gene: nuoJ; RP790

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 6

C;Keywords: membrane-associated complex; NAD; oxidoreductase

Query Match  
Best Local Similarity 45.8%; Score 44; DB 2; Length 205;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 5 NFNDVTTLREN 16  
77 NFNDVTTLREN 88

## RESULT 10

D69102

collagenase - Methanobacterium thermoautotrophicum (strain Delta H)

C;Species: Methanobacterium thermoautotrophicum

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999

R;Accession: D69102

R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

kl, S.; Church, G.M.; Daniels, C.J.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.

A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct

A;Reference number: A69000; MUID:98037514

A;Accession: D69102

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-807 <MTH>

A;Cross-references: GB:AE000931; GB:AE000666; MID:g2622885; PIDN:AAH86229.1; PID:g2622885

A;Experimental source: strain Delta H

A;Gene: MTH1763

A;Start codon: TTG

Query Match  
Best Local Similarity 45.8%; Score 44; DB 2; Length 807;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 1 FEHCNFDVTTLRE 15  
562 FECDMDVRSILRE 576

## RESULT 11

S41626

Spike protein chain 1 precursor - avian infectious bronchitis virus (fragment)

C;Species: avian infectious bronchitis virus, IBV

C;Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 20-Sep-1999

R;Accession: S41626

R;Cavanagh, D.; Davis, P.J.

A;Title: Sequence analysis of strains of avian infectious bronchitis coronavirus 1501

A;Reference number: S41626

A;Accession: S41626

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-544 <GAV>

A;Cross-references: EMBL:X64737; MID:9453158; PIDN:CAA46003.1; PID:9453159

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1992

C;Superfamily: coronavirus E2 glycoprotein

Query Match  
Best Local Similarity 44.8%; Score 43; DB 2; Length 544;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 3 HCNDVT 10  
101 HCNDVT 108

## RESULT 12

T20652

hypothetical protein Y102A5C.1 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999

R;Accession: T20652; T26370

R;Mortimore, B.

A;Title: The EMBL Data Library, November 1996

A;Reference number: T20652

A;Accession: T20652

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-365 <MIL>

A;Cross-references: EMBL:281496; PIDN:CAB04075.1; GSPDB:GN00023; CESP:Y102A5C.1

A;Experimental source: clone F09C6

R;Gardner, A.

A;Title: The EMBL Data Library, September 1998

A;Reference number: Z20204

A;Accession: T26370

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-365 <MIL>

A;Cross-references: EMBL:AL031627; PIDN:CAA20972.1; CESP:Y102A5C.1

A;Experimental source: clone Y102A5C

A;Gene: CESP:Y102A5C.1

A;Map position: 5

A;Introns: 93/2; 104/1; 298/3

C:Date: 01-Dec-2000 #sasquence\_revision 01-Dec-2000 #text\_change 31-Dec-2000  
C:Accession: C83681  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirata,  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and  
A:Reference number: AB3650; MUID:20263314  
A:Accession: C83681  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-419 <STO>  
A:Cross-references: GB:AP001507; GB:BA000004; MTD:q10172612; PIDN:BAB03970.1; GSPDB:GN000  
C:Genetics:  
A:Experimental source: strain C-125  
A:Gene: BH0251

Query Match	43.8%	Score 42;	DB 2;	Length 419;
Best Local Similarity	40.0%	Pred. No. 37;		
Matches	6;	Conservative	6;	Mismatches
			3;	Indels
			0;	Gaps
0Y	3	HCNFNDVTTLRLENE	17	
		: : : : : :		
Db	219	HTDYNTIISRIESE	233	

RESULT 14  
T16889  
hypothetical protein T19C3.5 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T16889  
R:Fulton, L.  
submitted to the EMBL Data Library, June 1995  
A:Description: The sequence of C. elegans cosmid T19C3.  
A:Reference number: Z18598  
A:Accession: T16889  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-464 <FULL>  
A:Cross-references: EMBL:U28412; NID:g849235; PID:g849239; PIDN:AAC46596.1; CESP:T19C3.5  
A:Experimental source: strain Bristol N2  
C:Genetics:  
A:Gene: CESP:T19C3.5  
A:Introns: 42/3; 82/2; 116/3; 195/2; 223/1; 333/3; 360/3; 407/2

Query Match	43.8%	Score 42;	DB 2;	Length 464;
Best Local Similarity	50.0%	Pred. No. 41;		
Matches	7;	Conservative	3;	Mismatches 4;
				Indels 0;
				Gaps 0;
QY	3	HCNFNDVTRLRKN	16	
DB	108	HAKEVEVITRLKRN	121	

RESULT 15  
C65969  
hypothetical protein yqzD - *Bacillus subtilis*  
C.Species: *Bacillus subtilis*  
C.Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 28-Jul-2000  
C.Accession: C65969  
R.Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berta  
C.; Bron, S.; Bouilliet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chn  
A.; Erlich, S.D.; Emmertson, P.T.; Estlin, K.D.; Eyrington, J.; Fabret, C.; Ferrari, E.  
Nature 350, 249-256, 1997  
A.Authors: Foulter, D.; Fittz, C.; Fujita, M.; Fujita, Y.; Funai, S.; Gallizzi, A.; Gall  
isch, J.; Hatwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Huilo, M.F.  
Joettler, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
A.Authors: Lamber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mene  
Y.M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portelle  
Rieger, M.; Rivolta, C.; Kocha, E.; Koche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanno  
A.Authors: Schleich, S.; Schroeter, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
A.Kuchl, M.; Tanakoshi, A.; Tanaka, T.; Terpiltra, P.; Tognoni, A.; Tosato, V.; Uchiyama

T.: Winters, P. Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida  
A.: Authors: Yoshikawa, H.F.; Zunshtein, E.; Yoshikawa, H.; Danchin, A.  
A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*  
A:Reference number: A69580; MUID:98044033  
A:Accession: C69969  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-117 <KUN>  
A:Cross-references: GB:Z99116; GB:AL009126; NID:g2634723; PID:g26349  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yqzd  
C:Superfamily: *Bacillus subtilis* hypothetical protein yqzd

```

Query Match Score 41; DB 2; Length 117;
Best Local Similarity 43.8%;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 2 EHCNENDVTRLRENE 17
   || : || : || : || : ||
Db 98 EHVSVEDVTLIRKNE 113

```

RESULT 16  
S23358  
H+-transporting ATP synthase (EC 3.6.1.34) chain b - Odontella sinensis chloroplast  
N:Alternate names: ATP synthase CF0 chain I; H+-transporting ATP synthase chain I  
C:Species: chloroplast Odontella sinensis  
C:Date: 03-May-1994 #sequence\_revision 20-Feb-1995 #text\_change 20-Jun-2000  
C:Accession: S23358; S14444; S78319  
J: Mol. Biol. 224, 529-536, 1992  
A:Title: Chloroplast ATPase genes in the diatom Odontella sinensis reflect cyanobacterial  
A:Reference number: S21682; MUID:92219274  
A:Accession: S23358  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-179 <PAN1>  
A:Cross-references: EMBL:X60752; NID:q11937; PIDN:CAA43162.1; PID:q11942  
R:Pancic, P.G.; Strothmann, H.; Kowallik, K.V.  
PDBS Lett. 280, 387-392, 1991  
A:Title: The delta subunit of the chloroplast ATPase is plastid-encoded in the diatom  
A:Reference number: S14323; MUID:91192176  
A:Accession: S14444  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 141-179 <PAN2>  
A:Cross-references: GB:X57701; NID:q14183  
A:Note: this ORF is not annotated in Genbank entry PLOSATPD, release 109.0  
R:Kowallik, K.V.; Stoebe, B.; Schafrin, I.; Kroth-Pancic, P.; Freier, U.  
Plant Mol. Biol. Rep. 13, 336-342, 1995  
A:Title: The Chloroplast Genome of a Chlorophyll a+c- containing Alga, Odontella sinensis  
A:Reference number: S78328  
A:Accession: S78319  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-179 <KOW>  
A:Cross-references: EMBL:Z67753; NID:q1185127; PIDN:CAA91692.1; PID:q1185209  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995  
C:Genetics:  
A:Gene: atpF  
A:Genome: chloroplast  
C:Superfamily: H+-transporting ATP synthase chain I  
C:Keywords: ATP biosynthesis; chloroplast; hydrolase; membrane-associated complex; thylakoid

```

Query Match Similarity      42.7%;   Score 41:   DB 2:   Length 179;
Best Local Similarity      61.5%;   Pred. No. 22;
Matches      8;   Conservative      1;   Mismatches      4;   Indels      0;   Gaps      0.
QY      5  NNNDVTRRLRENE 17
      ||| : | | |||

```

Db 3 NFNQITFLAENE 15

## RESULT 17

VCBVC

coat protein - tobacco rattle virus (strain CAM)

C:Species: tobacco rattle virus, TRV

C:Date: 30-Sep-1991

C:Accession: A26027 #sequence\_revision 30-Sep-1991 #text\_change 23-Jul-1999

R:Bergh, S.T.;

Nucleic Acids Res. 13, 8507-8518, 1985

A:Title: The nucleotide sequence of tobacco rattle virus RNA-2 (CAM strain).

A:Reference number: A26027; MUID:86093657

A:Accession: A26027

A:Molecule type: genomic RNA

A:Residues: 1-223 &lt;BER&gt;

A:Cross-references: GB:X03241; NID:962057; PIDN:CAA26998.1; PID:662058

C:Genetics:

A:Map position: segment 2

C:Superfamily: tobacco rattle virus coat protein

C:Keywords: coat protein

Query Match

Best Local Similarity 42.7%; Score 41; DB 1; Length 223;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 5 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14

1: |||||

25 NFNQITFLAENE 14



Query Match 42.7%; Score 41; DB 2; Length 742;  
Best Local Similarity 61.5%; Pred. No. 98;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CKNFNDVTRLREN 16  
| | | | |  
Db 384 CKINDEYTLREN 396

## RESULT 22

T03896  
hypothetical protein R09E12.6 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 01-Dec-2000  
C:Accession: T03896; T32395; T33999  
R:Blanchard, M.; Elliott, G.; Keppler, D.; Smith, A.  
Submitted to the EMBL Data Library, July 1997  
A:Description: The sequence of *C. elegans* cosmid R09E12.  
A:Reference number: Z15130  
A:Accession: T03896  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-805 <BLA>  
A:Cross-references: EMBL:AF016422; NID:92291187; PIDN:AB65317.1; PID:92291193  
R:Magri, L.; Scheet, P.  
Submitted to the EMBL Data Library, September 1997  
A:Description: The sequence of *C. elegans* cosmid K02E7.  
A:Reference number: Z21159  
A:Accession: T32395  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-281, 'H', 283-805 <MAG>  
A:Cross-references: EMBL:AF025465; PIDN:AA71023.1; GSPDB:GN00020; CESP:K02E7.3  
A:Experimental source: Strain Bristol N2; clone K02E7  
R:Paulley, A.; Scheet, P.; Harper, M.  
Submitted to the EMBL Data Library, February 1999  
A:Description: The sequence of *C. elegans* cosmid W03G1.  
A:Reference number: Z21454  
A:Accession: T33999  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-504, 'S', 506-805 <PAU>  
A:Cross-references: EMBL:AF123964; PIDN:AD14752.1; GSPDB:GN00022; CESP:W03G1.4  
A:Experimental source: strain Bristol N2; clone W03G1  
C:Genetics:  
A:Gene: CESP:K02E7.3; CESP:W03G1.4  
A:Map position: V  
A:Introns: 31/1; 196/2; 281/2; 357/3; 393/1; 429/3; 489/3; 533/3; 559/3; 609/2  
A:Note: R09E12.6  
C:Superfamily: *Caenorhabditis elegans* hypothetical protein W03G1.4

Query Match 42.7%; Score 41; DB 2; Length 805;  
Best Local Similarity 61.5%; Pred. No. 1.1e+02;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CKNFNDVTRLREN 16  
| | | | |  
Db 384 CKINDEYTLREN 396

## RESULT 23

T48898  
disease resistance protein RPP8 [validated] - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 21-Jul-2000  
C:Accession: T48898  
R:McDowell, J.M.; Dhandaydham, M.; Long, T.A.; Aarts, M.G.; Goff, S.; Holub, E.B.; Dangl  
Plant Cell 10, 1861-1874, 1998  
A:Title: Intragenic recombination and diversifying selection contribute to the evolution  
A:Reference number: Z24999; MUID:99030193  
A:Accession: T48898

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-906 <MCD>  
A:Cross-references: EMBL:AF089710; NID:93928861; PIDN:AA63165.1; PID:93928862  
A:Experimental source: Landsberg erecta  
C:Genetics:  
A:Gene: RPP8  
A:Introns: 293/1; 342/1  
A:Function:  
A:Description: promotes resistance to *Peronospora parasitica*

QY 2 EHCNFDVTRLRE 15  
| | | | |  
Db 703 ERCNFTLSLSRE 716

Query Match 42.7%; Score 41; DB 2; Length 906;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 EHCNFDVTRLRE 15  
| | | | |  
Db 703 ERCNFTLSLSRE 716

## RESULT 24

T48899  
disease resistance protein rpp8 [similarity] - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 21-Jul-2000  
C:Accession: T48899  
R:McDowell, J.M.; Dhandaydham, M.; Long, T.A.; Aarts, M.G.; Goff, S.; Holub, E.B.; Da  
Plant Cell 10, 1861-1874, 1998  
A:Title: Intragenic recombination and diversifying selection contribute to the evol  
A:Reference number: Z24999; MUID:99030193  
A:Accession: T48899  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-908 <MCD>  
A:Cross-references: EMBL:AF089711; NID:93901293; PIDN:AA678631.1; PID:93901294  
A:Experimental source: Columbia  
C:Genetics:  
A:Gene: rpp8  
A:Introns: 293/1; 342/1  
A:Function:  
A:Description: susceptible allele of a gene that promotes resistance to *Peronospora p*

Query Match 42.7%; Score 41; DB 2; Length 908;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 EHCNFDVTRLRE 15  
| | | | |  
Db 705 ERCNFTLSLSRE 718

## RESULT 25

S07421  
E2 glycoprotein precursor - avian infectious bronchitis virus  
N:Alternate names: Deploimer glycoprotein  
N:Contains: E2 glycoprotein subunit S1; E2 glycoprotein subunit S2  
C:Species: avian infectious bronchitis virus, IBV  
C>Date: 18-Feb-1994 #sequence\_revision 01-Dec-1995 #text\_change 20-Sep-1999  
C:Accession: S07421  
R:Niesters, H.G.M.; Lenstra, J.A.; Spaan, W.J.M.; Zijderfeld, A.J.; Bleumink-Pluym, N  
Virus Res. 5, 253-263, 1986  
A:Title: The deploimer protein sequence of the M1 strain of coronavirus IBV and its c  
A:Reference number: S07421; MUID:87021475  
A:Accession: S07421  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1162 <NIE>  
A:Cross-references: EMBL:M21883; NID:9331183; PIDN:AA66575.1; PID:9331184  
C:Superfamily: coronavirus E2 glycoprotein  
C:Keywords: glycoprotein

Tue Apr 2 09:51:57 2002

us-09-020-393b-3\_copy\_42\_58.open.rpr

Query Match  
Best Local Similarity 42.7%; Score 41; DB 2; Length 1162;  
Matches 6; Conservatively 75.0%; Pred. No. 1.6e+02;  
QY 3 HCNFNDVT 10 1; Mismatches 1; Indels 0; Gaps 0;  
Db 101 HCNFSDT 108

Search completed: April 2, 2002, 09:20:13  
Job time: 22 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:20:16 ; Search time 10.15 seconds  
(without alignments)  
61.409 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58  
Perfect score: 96  
Sequence: 1 FEHCNFNDVTRLRENE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	128	1	CD59_HUMAN
2	63	65.6	126	1	CD59_PAPSP
3	63	65.6	128	1	CD59_CERAE
4	49	51.0	124	1	CD59_RABIT
5	49	51.0	128	1	CD59_AOTTR
6	47	49.0	128	1	CD59_CALSO
7	42	43.8	464	1	Y5V5_CAEEL
8	41	42.7	179	1	ATPF_ODOSI
9	41	42.7	223	1	COAT_TKPCA
10	41	42.7	720	1	YD71_SCHPO
11	41	42.7	1162	1	VG12_IBVB
12	41	42.7	1162	1	VG12_IBVM
13	41	42.7	1528	1	KEM1_YEAST
14	41	42.7	1648	1	RKPO_CGMSV
15	41	42.7	3433	1	UTRO_HUMAN
16	40	41.7	165	1	RIMM_RICPR
17	40	41.7	212	1	HFA3_HAEIN
18	40	41.7	231	1	FLPA_SULAC
19	40	41.7	607	1	KMS1_SCHPO
20	40	41.7	716	1	HSB3_DROAV
21	40	41.7	742	1	SVY2_RAT
22	39.5	41.1	169	1	Y594_METJA
23	39.5	41.1	760	1	SPOT_SYNY3
24	39.5	41.1	854	1	ENV_STVCZ
25	39	40.6	123	1	CD59_PIG
26	39	40.6	224	1	PVA_PLAFA
27	39	40.6	289	1	Y084_MYCPN
28	39	40.6	297	1	STX4_HUMAN
29	39	40.6	353	1	YCS4_HAEIN
30	39	40.6	574	1	VB18_VACC
31	39	40.6	574	1	VB18_VACCV
32	39	40.6	688	1	YJ80_YEAST
33	39	40.6	1678	1	CA64_HUMAN

## ALIGNMENTS

RESULT	1	CD59_HUMAN	STANDARD;	PRT;	128 AA.
ID	CD59_HUMAN				
AC	P13987;				
DT	01-JAN-1990 (Rel. 13, Created)				
DT	01-JAN-1990 (Rel. 13, Last sequence update)				
DT	20-AUG-2001 (Rel. 40, Last annotation update)				
DE	CD59 GLYCOPROTEIN PRECURSOR (MEMBRANE ATTACK COMPLEX INHIBITION FACTOR) (MACIF) (MAC-INHIBITORY PROTEIN) (MAC-IP) (MEM43 ANTIGEN) (PROTECTIN) (MEMBRANE INHIBITOR OF REACTIVE LYSIS) (MIRL) (20 KDA DE HOMOLOGOUS RESTRICTION FACTOR) (HRR-20) (HRR20) (1F5 ANTIGEN) CD59.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=T-cell;				
RX	MEDLINE=89361238; PubMed=2475570;				
RA	Davies A., Simmons D.L., Hale G., Harrison R.A., Tighe H.,				
RA	Lachmann P.J., Waldmann H.;				
RT	"CD59, an IX-6-like protein expressed in human lymphoid cells,				
RT	regulates the action of the complement membrane attack complex on				
RT	homologous cells.";				
RL	J. Exp. Med. 170:637-654(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90168959; PubMed=1689664;				
RA	Philbrick W.M., Palfrey R.G.E., Roger G.E., Maher S.E.,				
RT	"The CD59 antigen is a structural homologue of murine Ly-6 antigens				
RT	but lacks interferon inducibility.";				
RL	Eur. J. Immunol. 20:87-92(1990).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89350983; PubMed=2475111;				
RA	Okada H., Nagami Y., Takahashi K., Okada N., Hideshima T.,				
RA	Takikawa H., Kondo J.;				
RT	"20 KDa homologous restriction factor of complement resembles T cell				
RT	activating protein.";				
RL	Biochem. Biophys. Res. Commun. 162:1553-1559(1989).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90110046; PubMed=2606909;				
RA	Sugita Y., Tobe T., Oda E., Tomita M., Yasukawa K., Yamaji N.,				
RA	Takemoto T., Furuchi K., Takayama M., Yano S.;				
RT	"Molecular cloning and characterization of MACIF, an inhibitor of				
RT	membrane channel formation of complement.";				
RL	J. Biochem. 106:555-557(1989).				
RN	[5]				

34	39	40.6	1888	1	YD72_SCHPO	014207 schizosacch
35	38	39.6	344	1	YD76_SCHPO	014220 schizosacch
36	38	39.6	502	1	KDC2_DROME	P16922 drosophila
37	38	39.6	1233	1	HCV4_OCTDO	P12659 octopus dof
38	38	39.6	1450	1	RPO1_ASEB7	P42486 african swi
39	38	39.6	2896	1	HCV6_OCTDO	061363 octopus dof
40	37.5	39.1	517	1	Y4CC_RH1SN	P53385 rhizobium s
41	37	38.5	137	1	COF1_DICDI	P54706 dictyostell
42	37	38.5	273	1	RL2_RICPR	092c98 rickettsia
43	37	38.5	420	1	PEL_BACSU	P39116 bacillus su
44	37	38.5	516	1	SMP3_YEAST	004174 saccharomyc
45	37	38.5	535	1	ATP2_YEAST	P53296 saccharomyc
46	37	38.5	684	1	YMW5_YEAST	004779 saccharomyc
47	37	38.5	758	1	PURL_CORAM	091hw9 corynebacte
48	37	38.5	853	1	VMTH_LAMBD	P03736 bacterioph
49	37	38.5	875	1	CND3_SCHPO	010429 schizosacch
50	37	38.5	1082	1	SP23_YEAST	P35210 saccharomyc

RP SEQUENCE FROM N.A.  
RX MEDLINE-90253615; PubMed-1692709;  
RA Sawada R., Ohashi K., Anaguchi H., Okazaki H., Hattori M., Minato N.,  
RT Naruto M.;  
RT "Isolation and expression of the full-length cDNA encoding CD59  
RL antigen of human lymphocytes.";  
RN DNA Cell Biol. 9:213-220(1990).  
[6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-92390353; PubMed-1381503;  
RA Petranka J.G., Fleener D.E., Sykes K., Kaufman R.E., Rosse W.F.;  
RT "Structure of the CD59-encoding gene: further evidence of a  
RT relationship to murine lymphocyte antigen Ly-6 protein.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:7876-7879(1992).  
[7]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Blood;  
RX MEDLINE-93021133; PubMed-1383553;  
RA Tone M., Walsh L.A., Waldmann H.;  
RT "Gene structure of human CD59 and demonstration that discrete mRNAs  
RT are generated by alternative polyadenylation.";  
RL J. Mol. Biol. 227:971-976(1992).  
[8]  
RP SEQUENCE OF 27-128 FROM N.A.  
RX MEDLINE-89386002; PubMed-2476718;  
RA Sawada R., Ohashi K., Okano K., Hattori M., Minato N., Naruto M.;  
RT "Complementary DNA sequence and deduced peptide sequence for  
RT CD59/MEM-43 antigen, the human homologue of murine lymphocyte antigen  
RL Ly-6C.";  
RN Nucleic Acids Res. 17:6728-6728(1989).  
[9]  
RP GPI-ANCHOR AND DISULFIDE BONDS.  
RX MEDLINE-94103166; PubMed-8276756;  
RA Sugita Y., Nakano Y., Oda E., Noda K., Tobe T., Miura N.H., Tomita M.;  
RT "Determination of carboxyl-terminal residue and disulfide bonds of  
RT MACP (CD59), a glycosyl-phosphatidylinositol-anchored membrane  
RL protein.";  
RN J. Biochem. 114:473-477(1993).  
[10]  
RP BINDING STUDIES.  
RX MEDLINE-92317101; PubMed-1377690;  
RA Nishimura H., Sims P.J.;  
RT "The human complement regulatory protein CD59 binds to the alpha-  
RL chain of C8 and to the 'b' domain of C9.";  
RN J. Biol. Chem. 267:13675-13680(1992).  
[11]  
RP MUTATIONAL ANALYSIS.  
RX MEDLINE-97188784; PubMed-9053451;  
RA Bodian D.L., Davis S.J., Morgan B.P., Rushmere N.K.;  
RT "Mutational analysis of the active site and antibody epitopes of the  
RL complement-inhibitory glycoprotein, CD59.";  
RN J. Exp. Med. 185:507-516(1997).  
[12]  
RP IDENTIFICATION OF COMPLEMENT INHIBITORY DOMAIN.  
RX MEDLINE-97383147; PubMed-9235986;  
RA Yu J., Dong S., Rushmere N.K., Morgan B.P., Abegyan R., Tomlinson S.;  
RT "Mapping the regions of the complement inhibitor CD59 responsible for  
RL its species selective activity.";  
RN Biochemistry 36:9423-9428(1997).  
[13]  
RP STRUCTURE OF CARBOHYDRATES AND GPI-ANCHOR, AND SEQUENCE OF N-TERMINUS.  
RX MEDLINE-97207284; PubMed-9054419;  
RA Rudd P.M., Morgan B.P., Wormald M.R., Harvey D.J., van den Berg C.W.,  
RT "The glycosylation of the complement regulatory protein, human  
RL erythrocyte CD59.";  
RN J. Biol. Chem. 272:7229-7244(1997).  
[14]  
RP INHIBITION BY GLYCATION, AND MUTAGENESIS OF LYS-66 AND HIS-69.  
RX MEDLINE-20266386; PubMed-10805801;  
RA Acosta J., Hettler J., Flueckiger R., Krumrei N., Goldfine A.,  
RT "Molecular basis for a link between complement and the vascular  
RN complications of diabetes.";  
RN Proc. Natl. Acad. Sci. U.S.A. 97:5450-5455(2000).  
[15]  
RP STRUCTURE BY NMR OF 26-95.  
RX MEDLINE-94213818; PubMed-7512825;  
RA Kiefer B., Driscoll P.C., Campbell I.D., Willis A.C.,  
RT van der Merwe P.A., Davis S.J.;  
RT "Three-dimensional solution structure of the extracellular region of  
RL the complement regulatory protein CD59, a new cell-surface protein  
RN domain related to snake venom neurotoxins.";  
RN Biochemistry 33:4471-4482(1994).  
[16]  
RP STRUCTURE BY NMR OF 26-102.  
RC TISSUE-Urine;  
RX MEDLINE-94348877; PubMed-7520819;  
RA Fletcher C.M., Harrison R.A., Lachmann P.J., Neuhaus D.;  
RT "Structure of a soluble, glycosylated form of the human complement  
RL regulatory protein CD59.";  
RN Structure 2:185-199(1994).  
-1- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK  
COMPLEX (MAC) ACTION. ACTS BY BINDING TO THE C8 AND/OR C9  
INCORPORATION OF THE ASSEMBLING MAC, THEREBY PREVENTING  
FORMATION OF THE OSMOTIC PORE. THIS INHIBITOR APPEARS TO BE  
SPECIES-SPECIFIC. INVOLVED IN SIGNAL TRANSDUCTION FOR T-CELL  
ACTIVATION COMPLEXED TO A PROTEIN TYROSINE KINASE. INTERACTS WITH  
T-CELL SURFACE ANTIGEN CD2.  
-1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.  
-1- PTM: N-AND O-GLYCOSYLATED. THE N-GLYCOSYLATION MAINLY CONSISTS OF A  
FAMILY OF BI-ANTENNARY COMPLEX-TYPE STRUCTURES WITH AND WITHOUT  
LACTOSAMINE EXTENSIONS AND OUTER ARM FUCCOSE RESIDUES. THE  
PREDOMINANT O-GLYCANS ARE MONO-SIALYLATED FORMS OF THE  
DISACCHARIDE, GAL-BETA-1,3GALNAc, AND THEIR SITES OF ATTACHMENT  
ARE PROBABLY ON THR-76 AND THR-77.  
-1- PTM: GLYCATED. GLYCATION IS FOUND IN DIABETIC SUBJECTS, BUT ONLY  
AT MINIMAL LEVELS IN NONDIABETIC SUBJECTS. GLYCATED CD59 LACKS  
MAC-INHIBITORY FUNCTION AND CONFERS TO VASCULAR COMPLICATIONS OF  
DIABETES.  
-1- SIMILARITY: CONTAINS 1 UPAR/LY6 DOMAIN.  
-1- DATABASE: NAME-PROW; NOTE-CD guide CD59 entry;  
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd59.htm".  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
-----  
DR EMBL; M27909; AAA60543.1; -  
DR EMBL; M95708; AAA60957.1; -  
DR EMBL; X16447; CAA34467.1; -  
DR EMBL; X17198; CAA35059.1; -  
DR EMBL; X15861; CAA33870.1; -  
DR EMBL; M34671; AAA51952.1; -  
DR EMBL; M84345; - NOT\_ANNOTATED\_CDS.  
DR EMBL; M84349; AAA88793.1; -  
DR EMBL; M84346; AAA88793.1; JOINED.  
DR EMBL; M84348; AAA88793.1; JOINED.  
DR EMBL; Z14113; CAA78486.1; -  
DR EMBL; Z14114; CAA78486.1; JOINED.  
DR EMBL; Z14115; CAA78486.1; JOINED.  
DR PIR; J10109; RWH059.  
DR PIR; A34587; A34587.  
DR PIR; S05504; S05504.  
DR PIR; S09201; S09201.  
DR PIR; A46252; A46252.  
DR PDB; 1CDQ; 30-SEP-94.  
DR PDB; 1CDR; 30-SEP-94.  
DR PDB; 1CDG; 30-SEP-94.  
DR PDB; 1ERG; 30-APR-94.  
DR PDB; 1ERH; 30-APR-94.

DR GLYcosuiteDB; P13987; -  
DR MIM; 107271; -  
DR InterPro: IPR001526; LY6\_UPAR.  
DR Pfam: PF00021; UPAR\_LY6; 1.  
DR SMART; SM00134; LU; 1.

Query Match 100.0%; Score 96; DB 1; Length 128;  
Best Local Similarity 100.0%; Pred. No. 2.8e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHCNNDVTTLRRE 17  
Db 67 FEHCNNDVTTLRRE 83

## RESULT 2

CD59\_PAPSP STANDARD; PRT; 126 AA.  
AC Q28785;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE CD59 GLYCOPROTEIN PRECURSOR (MEMBRANE ATTACK COMPLEX INHIBITION  
DE FACTOR) (MACIF) (MAC-INHIBITORY PROTEIN) (MAC-IP) (PROTECTIN).  
GN CD59.  
OS Papio sp. (Baboon).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Papio.  
OX NCBI\_TaxId=61183;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95104908; PubMed=7528724;  
RA Fodor W.L., Rollins S.A., Bianco-Caron S., Burton W.V.,  
RA Guilmette E.R., Rother R.P., Zavolico G.B., Squitto S.P.,  
RT "Primate terminal complement inhibitor homologues of human CD59.";  
RL Immunogenetics 41:51-51(1995).  
CC -1- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK  
CC COMPLEX (MAC) ACTION. ACTS AT OR AFTER THE C5B-8 STAGE OF MAC  
CC ASSEMBLY.  
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR  
CC (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 1 UPAR/LY6 DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; L22862; AAA74127.1; -  
DR HSSP; P13987; ICDO.  
DR InterPro: IPR001526; LY6\_UPAR.  
DR Pfam; PF00021; UPAR\_LY6; 1.  
DR SMART; SM00134; LU; 1.  
DR PROSITE; PS00983; LY6\_UPAR; 1.  
KW Antigen; Glycoprotein; GPI-anchor; Signal.  
FT SIGNAL 1 25  
FT CHAIN 1 100  
FT PROPEP 101 126  
FT DOMAIN 26 106  
FT DISULFID 28 51  
FT DISULFID 31 38  
FT DISULFID 44 64  
FT DISULFID 70 88  
FT DISULFID 89 94  
FT CARBOHYD 43 43  
FT LIPID 100 100  
SQ SEQUENCE 126 AA; 13716 MW; 7900FF937871EBDC CRC64;

Query Match 65.6%; Score 63; DB 1; Length 126;  
Best Local Similarity 58.8%; Pred. No. 0.00095;  
Matches 10; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 FEHCNNDVTTLRRE 17  
Db 67 FANCFNDISTLLKESE 83

## RESULT 3

CD59\_CERAE STANDARD; PRT; 128 AA.  
ID CD59\_CERAE  
AC Q28216;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE CD59 GLYCOPROTEIN PRECURSOR (MEMBRANE ATTACK COMPLEX INHIBITION  
DE FACTOR) (MACIF) (MAC-INHIBITORY PROTEIN) (MAC-IP) (PROTECTIN).  
GN CD59.  
OS Cercopithecus aethiops (Green monkey) (Girvet).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Cercopithecus.  
OX NCBI\_TaxId=9534;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95104908; PubMed=7528724;  
RA Fodor W.L., Rollins S.A., Bianco-Caron S., Burton W.V.,  
RA Guilmette E.R., Rother R.P., Zavolico G.B., Squitto S.P.,  
RT "Primate terminal complement inhibitor homologues of human CD59.";  
RL Immunogenetics 41:51-51(1995).  
CC -1- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK  
CC COMPLEX (MAC) ACTION. ACTS AT OR AFTER THE C5B-8 STAGE OF MAC  
CC ASSEMBLY.  
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR  
CC (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 1 UPAR/LY6 DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; L22863; AAA74126.1; -  
DR HSSP; P13987; ICDO.  
DR InterPro: IPR001526; LY6\_UPAR.  
DR Pfam; PF00021; UPAR\_LY6; 1.  
DR SMART; SM00134; LU; 1.  
DR PROSITE; PS00983; LY6\_UPAR; 1.  
KW Antigen; Glycoprotein; GPI-anchor; Signal.  
FT SIGNAL 1 25  
FT CHAIN 1 102  
FT PROPEP 103 128  
FT DOMAIN 102 102  
FT DISULFID 26 108  
FT DISULFID 28 51  
FT DISULFID 31 38  
FT DISULFID 44 64  
FT DISULFID 70 88  
FT DISULFID 89 94  
FT CARBOHYD 43 43  
SQ SEQUENCE 128 AA; 14007 MW; 9778DEF7F05152 CRC64;

Query Match 65.6%; Score 63; DB 1; Length 128;  
Best Local Similarity 58.8%; Pred. No. 0.00096;  
Matches 10; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 FEHCNNDVTTLRRE 17  
Db 67 FANCFNDISTLLKESE 83

Db 67 FANCFNDISTLKRESE 83

```
RESULT 4
CD59_RABIT
ID CD59_RABIT STANDARD: PRT: 124 AA.
AC 077541;
DT 20-AUG-2001 (Rel. 40, Created)
DE 20-AUG-2001 (Rel. 40, last sequence update)
DE CD59_GLYCOPROTEIN PRECURSOR (MEMBRANE ATTACK COMPLEX INHIBITION
DE FACTOR) (MACIF) (MAC-INHIBITORY PROTEIN) (MAC-IP) (PROTECTIN).
GN CD59.
OC Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN NCBI_TaxID=9986;
RP SEQUENCE FROM N.A. AND SEQUENCE OF 25-64.
RX MEDLINE=98221206; PubMed=955129;
RA Zhao X.-J., Zhou J., Zhou Q., Sims P.J.;
RT "Identity of the residues responsible for the species-restricted
RT complement inhibitory function of human CD59."
RL J. Biol. Chem. 273:10665-10671(1998).
CC -1- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK
CC COMPLEX (MAC) ACTION. ACTS BY BINDING TO THE C8 AND/OR C9
CC INCOMPORATION OF THE ASSEMBLING MAC, THEREBY PREVENTING
CC FORMATION OF THE MULTIPLE COPIES OF C9 REQUIRED FOR COMPLETE
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
CC (BY SIMILARITY).
CC -1- MISCELLANEOUS: THE MATURE FORM OF THIS CD59 CONTAINS AN ADDITIONAL
CC SERINE RESIDUE BEFORE THE CONSERVED N-TERMINAL LEUCINE RESIDUE
CC FOUND IN ALL OTHER CD59 HOMOLOGS SEQUENCED TO DATE.
CC -1- SIMILARITY: CONTAINS 1 UPAR/LY6 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF040387; AAC23590.1; -
DR HSSP: P13987; 1ERG.
DR InterPro: IPR001526; LY6_UPAR.
DR Pfam: PF00021; UPAR_LY6; 1.
DR SMART: SM00134; LY; 1.
DR PROSITE: PS00983; LY6_UPAR; FALSE_NEG.
DR Antigen; Glycoprotein; GPI-anchor; signal.
FT SIGNAL 1 24
FT CHAIN 1 24
FT PROPEP 25 101 CD59_GLYCOPROTEIN.
FT DOMAIN 102 124 REMOVED IN MATURE FORM (BY SIMILARITY).
FT DISULFID 25 101 UPAR/LY6.
FT DISULFID 28 51 BY SIMILARITY.
FT DISULFID 31 38 BY SIMILARITY.
FT DISULFID 44 64 BY SIMILARITY.
FT DISULFID 70 88 BY SIMILARITY.
FT DISULFID 89 94 BY SIMILARITY.
FT CARBOHYD 37 37 N-LINKED (GLCNAC... ) (POTENTIAL).
FT LIPID 101 101 GPI-ANCHOR (BY SIMILARITY).
SQ SEQUENCE 124 AA; 13870 MW; 6EA64C816772CABD CRC64;
```

Query Match 51.0%; Score 49; DB 1; Length 124;  
Best Local Similarity 50.0%; Pred. No. 0.21;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 FEHCNFNDVTTLRLEN 16  
Db 67 YEDCNFEFISNRLEN 82

```
RESULT 5
CD59_AOTTR
ID CD59_AOTTR STANDARD: PRT: 128 AA.
AC P51447;
DT 01-OCT-1996 (Rel. 34, Created)
DE 01-OCT-1996 (Rel. 34, last sequence update)
DE 01-NOV-1997 (Rel. 35, last annotation update)
DE CD59_GLYCOPROTEIN PRECURSOR (MEMBRANE ATTACK COMPLEX INHIBITION
DE FACTOR) (MACIF) (MAC-INHIBITORY PROTEIN) (MAC-IP) (PROTECTIN).
GN CD59.
OC Aotus trivirgatus (Night monkey) (Douroucoulli).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
RN NCBI_TaxID=9505;
RP SEQUENCE FROM N.A.
RX MEDLINE=95104908; PubMed=7528724;
RA Fodor W.L., Rollins S.A., Blanco-Caron S., Burton W.V.,
RA Guilmette E.R., Rother R.P., Zavolco G.B., Squinto S.P.;
RT "Primate terminal complement inhibitor homologues of human CD59."
RL Immunogenetics 41:51-51(1995).
CC -1- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK
CC COMPLEX (MAC) ACTION. ACTS AT OR AFTER THE C5B-8 STAGE OF MAC
CC ASSEMBLY.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
CC (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 UPAR/LY6 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L22861; AAA35372.1; -
DR HSSP: P13987; 1CDO.
DR InterPro: IPR001526; LY6_UPAR.
DR Pfam: PF00021; UPAR_LY6; 1.
DR SMART: SM00134; LY; 1.
DR PROSITE: PS00983; LY6_UPAR; 1.
DR Antigen; Glycoprotein; GPI-anchor; signal.
FT SIGNAL 1 25
FT CHAIN 1 25
FT PROPEP 26 102 CD59_GLYCOPROTEIN.
FT DOMAIN 103 128 REMOVED IN MATURE FORM (BY SIMILARITY).
FT DISULFID 28 108 BY SIMILARITY.
FT DISULFID 31 38 BY SIMILARITY.
FT DISULFID 44 64 BY SIMILARITY.
FT DISULFID 70 88 BY SIMILARITY.
FT DISULFID 89 94 BY SIMILARITY.
FT CARBOHYD 43 43 N-LINKED (GLCNAC... ) (POTENTIAL).
FT LIPID 102 102 GPI-ANCHOR (BY SIMILARITY).
SQ SEQUENCE 128 AA; 14200 MW; 62D219B5589E55B CRC64;
```

Query Match 51.0%; Score 49; DB 1; Length 128;  
Best Local Similarity 52.9%; Pred. No. 0.22;  
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 FEHCNFNDVTTLRLEN 17  
Db 67 FEDCTFSRVSNQLSENE 83

RESULT 6  
CD59\_CALSO  
ID CD59\_CALSO STANDARD: PRT: 128 AA.  
AC P46657;  
DT 01-NOV-1995 (Rel. 32, Created)

```

DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CD59 GLYCOPROTEIN PRECURSOR (MEMBRANE ATTACK COMPLEX INHIBITION
DE FACTOR) (MACIF) (MAC-INHIBITORY PROTEIN) (MAC-IP) (PROTECTIN).
GN CD59.
OS Callitrix sp. (Marmoset).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC Callitrix.
OX NCBI_TaxId=9485;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95104908; PubMed=7528724;
RA Fodor W.L., Rollins S.A., Bianco-Caron S., Burton W.V.,
RA Guilmette E.R., Rother R.P., Zavoico G.B., Squitto S.P.;
RT "Primate terminal complement inhibitor homologues of human CD59."
RU Immunogenetics 41:51-51(1995).
CC -I- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK
CC COMPLEX (MAC) ACTION. ACTS AT OR AFTER THE C5B-8 STAGE OF MAC
CC ASSEMBLY.
CC -I- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
CC (BY SIMILARITY).
CC -I- SIMILARITY: CONAINS 1 UPAR/LY6 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L2860; AAA53398.1; .
DR HSSP: P13987; ICDO.
DR InterPro: IPR001526; LY6_UPAR.
DR Pfam: PF00021; UPAR_LY6; 1.
DR SMART: SM00134; LU; 1.
DR PROSITE: PS00983; LY6_UPAR; 1.
KW Antigen; Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 25 BY SIMILARITY.
FT CHAIN 26 102 CD59 GLYCOPROTEIN.
FT PROPE 103 128 REMOVED IN MATURE FORM (BY SIMILARITY).
FT DOMAIN 26 108 UPAR/LY6.
FT DISULFID 28 51 BY SIMILARITY.
FT DISULFID 31 38 BY SIMILARITY.
FT DISULFID 44 64 BY SIMILARITY.
FT DISULFID 70 88 BY SIMILARITY.
FT DISULFID 89 94 BY SIMILARITY.
FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 102 102 GPI-ANCHOR (BY SIMILARITY).
SQ SEQUENCE 128 AA; 14210 MW; 7A44CCAADECEDB4 CRC64;

Query Match 49.0%; Score 47; DB 1; Length 128;
Best Local Similarity 47.1%; Pred. No. 0.47;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 FHCNPNNDVTRRLRENE 17
Db 67 FEDCTFRQLSNOLSENE 83

RESULT 7
YVS5_CAEEL STANDARD; PRT; 464 AA.
AC Q10011;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOHETICAL 51.9 KDA PROTEIN T19C3.5 IN CHROMOSOME III PRECURSOR.
GN T19C3.5
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidae;

```

```

OC Rhabditidae; Peleoderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Fulton L.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: SOME, TO THE BPI/CEP/LBP/PLTP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U28412; AAC46596.1; .
DR WormPep: T19C3.5; CE02036.
DR InterPro: IPR001124; LBP_BPI_CETP.
DR Pfam: PF01273; LBP_BPI_CETP; 1.
DR SMART: SM00328; BPI1; 1.
DR SMART: SM00329; BPI2; 1.
DR PROSITE: PS00400; LBP_BPI_CETP; FALSE_NEG.
KW Hypothetical protein; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 464 HYPOHETICAL PROTEIN T19C3.5.
SQ SEQUENCE 464 AA; 51939 MW; 3029CFD2F65E67F CRC64;

Query Match 43.8%; Score 42; DB 1; Length 464;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 HCNPNNDVTRRLREN 16
Db 108 HAKETVYTRLRKN 121

RESULT 8
ATPF_ODOSI STANDARD; PRT; 179 AA.
ID ATPF_ODOSI
AC Q00822;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ATP SYNTHASE B CHAIN (EC 3.6.1.34) (SUBUNIT 1).
GN ATPF.
OS Odontella sinensis.
OC Chlorophyta.
OC Eukaryota; streptophytes; Bacillariophyta; Coscinodiscophyceae;
OC Biddulphiophycidae; Eupodiscaceae; Odontella.
OX NCBI_TaxId=2839;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92219274; PubMed=1532839;
RA Pantic P.G., Strotmann H., Kowallik K.V.;
RT "Chloroplast Arpase genes in the diatom Odontella sinensis reflect
RT cyanobacterial characters in structure and arrangement."
RL J. Mol. Biol. 224:529-536(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Kowallik K.V., Stoebe B., Schaffran I., Kroth-Pantic P., Freier U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga,
RT Odontella sinensis."
RN [3]
RP SEQUENCE OF 141-179 FROM N.A.
RX MEDLINE=91192176; PubMed=1826484;
RA Pantic P.G., Strotmann H., Kowallik K.V.;
RT "The delta subunit of the chloroplast Arpase is plastid-encoded in
RT the diatom Odontella sinensis."
RL FEBS Lett. 280:387-392(1991).

```

CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC  
CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE  
SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), Epsilon(1). CF(0)  
HAS THREE MAIN SUBUNITS: A, B AND C.  
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE.  
CC -1- SIMILARITY: BELONGS TO THE ATPASE B CHAIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: X60752; CAA43155.1; -  
CC EMBL: 267753; CAA91692.1; -  
CC EMBL: X57701; NOT\_ANNOTATED\_CDS.  
CC PIR: S14444; S14444.  
CC PIR: S23358; S23358.  
CC Mendel; 4844; ODOsi:atpf:1  
CC InterPro: IPR002146; ATP-synt\_B.  
CC Pfam: PF00430; ATP-synt\_B; 1.  
CC TRANSMEM 24  
CC Hydrogen ion transport; Transmembrane; CF(0); Chloroplast.  
CC PT  
CC SQ SEQUENCE 179 AA; 20074 MW; 4B614F93D501483D CRC64;  
  
Query Match  
Best Local Similarity 42.7%; Score 41; DB 1; Length 179;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
  
OY 5 NENDVTRLR 17  
Db 3 NFNQIFTLAENE 15  
||| : ||| |||  
-----  
RESULT 9  
COAT\_TYRCA STANDARD; PRT; 223 AA.  
ID AC P05070;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DE COAT PROTEIN (CAPSID PROTEIN).  
OS Tobacco rattle virus (strain CAM).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobravirus.  
OX NCBI\_TaxID=12296;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86093657; PubMed=3841203;  
RA Bergin S.T., Koziel M.G., Huang S.-C., Thomas R.A., Gilley D.P.,  
RT Siegel A.;  
RL "The nucleotide sequence of tobacco rattle virus RNA-2 (CAM strain).";  
Nucleic Acids Res. 13:8507-8518(1985).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: X03241; CAA26998.1; -  
CC PIR: A26027; VCBVCA.  
CC Coat protein.  
CC SQ SEQUENCE 223 AA; 23683 MW; EF1652A2F743398 CRC64;

Query Match  
Best Local Similarity 42.7%; Score 41; DB 1; Length 223;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 NENDVTRLR 14  
Db 25 NMDVTRLR 34  
||| : |||||  
-----  
RESULT 10  
ID YD71\_SCHPO STANDARD; PRT; 720 AA.  
AC Q10326;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DE HYPOTHETICAL 82.5 KDA PROTEIN C32A11.01 IN CHROMOSOME I.  
GN SPAC32A11.01.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetes; Schizosaccharomycetaceae;  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX SKELTON J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;  
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: Z69796; CAA93698.1; -  
CC KW Hypothetical protein.  
CC SQ SEQUENCE 720 AA; 82456 MW; D4DD9E6FC6E5604C CRC64;

Query Match  
Best Local Similarity 42.7%; Score 41; DB 1; Length 720;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
OY 1 FENCFNPDYTT 11  
Db 383 FSHCSFESLTT 393  
||| : ||| : ||  
-----  
RESULT 11  
ID VGL2\_IBVB STANDARD; PRT; 1162 AA.  
AC P11223; P05134;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLOMER PROTEIN)  
GN [CONAINS: SPIKE PROTEIN S1; SPIKE PROTEIN S2].  
OS Avian infectious bronchitis virus (strain Beaudette) (IBV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
OC Coronaviridae; Coronaviruses.  
OX NCBI\_TaxID=11122;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85159540; PubMed=2984314;  
RA Binn M.M., Boursnell M.E.G., Cavanagh D., Pappin D.J.C.,  
RT "Cloning and sequencing of the gene encoding the spike protein of the  
coronavirus IBV.";  
RL J. Gen. Virol. 66:719-726(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87085499; PubMed=3025348;

Query Match  
Best Local Similarity 42.7%; Score 41; DB 1; Length 720;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
OY 1 FENCFNPDYTT 11  
Db 383 FSHCSFESLTT 393  
||| : ||| : ||  
-----  
RESULT 11  
ID VGL2\_IBVB STANDARD; PRT; 1162 AA.  
AC P11223; P05134;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLOMER PROTEIN)  
GN [CONAINS: SPIKE PROTEIN S1; SPIKE PROTEIN S2].  
OS Avian infectious bronchitis virus (strain Beaudette) (IBV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
OC Coronaviridae; Coronaviruses.  
OX NCBI\_TaxID=11122;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85159540; PubMed=2984314;  
RA Binn M.M., Boursnell M.E.G., Cavanagh D., Pappin D.J.C.,  
RT "Cloning and sequencing of the gene encoding the spike protein of the  
coronavirus IBV.";  
RL J. Gen. Virol. 66:719-726(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87085499; PubMed=3025348;



RA Bins M.M., Boursnell M.E.G., Tomley F.M., Brown D.K.;  
 RT "Comparison of the spike precursor sequences of coronavirus IBV  
 RL strains M41 and 6/82 with that of IBV Beaudette."; J. Gen. Virol. 67:2825-2831(1986).  
 CC -I- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS  
 CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M95169; AAA70235.1; -  
 DR EMBL: X02342; CA26201.1; -  
 DR InterPro: IPR002551; Corona\_S1.  
 DR InterPro: IPR002552; Corona\_S2.  
 DR Pfam: PF01600; Corona\_S1; 1.  
 DR Pfam: PF01601; Corona\_S2; 1.  
 DR Glycoprotein; Envelope protein; Transmembrane; Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 1162  
 FT CHAIN 19 537  
 FT CHAIN 538 1162  
 FT DOMAIN 1120 1137  
 FT CARBOHYD 51 51  
 FT CARBOHYD 77 77  
 FT CARBOHYD 103 103  
 FT CARBOHYD 144 144  
 FT CARBOHYD 163 163  
 FT CARBOHYD 178 178  
 FT CARBOHYD 212 212  
 FT CARBOHYD 237 237  
 FT CARBOHYD 247 247  
 FT CARBOHYD 264 264  
 FT CARBOHYD 276 276  
 FT CARBOHYD 306 306  
 FT CARBOHYD 425 425  
 FT CARBOHYD 447 447  
 FT CARBOHYD 513 513  
 FT CARBOHYD 530 530  
 FT CARBOHYD 579 579  
 FT CARBOHYD 591 591  
 FT CARBOHYD 669 669  
 FT CARBOHYD 714 714  
 FT CARBOHYD 947 947  
 FT CARBOHYD 960 960  
 FT CARBOHYD 979 979  
 FT CARBOHYD 1014 1014  
 FT CARBOHYD 1038 1038  
 FT CARBOHYD 1051 1051  
 FT CARBOHYD 1074 1074  
 FT CARBOHYD 1162 1162  
 SQ SEQUENCE 1162 AA; 128046 MW; 0BAAD58113CEBBD5 CMC64;

Query Match 42.7%; Score 41; DB 1; Length 1162;  
 Best Local Similarity 75.0%; Pred. No. 50;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 HCNFNDVT 10  
 DB 101 HCNFSDTT 108

RESULT 12  
 VGL2\_IBVM STANDARD; PRT; 1162 AA.  
 AC P12651;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLIMER PROTEIN)  
 DE [CONTAINS: SPIKE PROTEIN S1; SPIKE PROTEIN S2].  
 GN S.  
 OS Avian infectious bronchitis virus (strain M41) (IBV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronaviruses.  
 OX NCBI\_TaxID=111127;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87021475; PubMed=2429473;  
 RA Niesters H.G.M., tenstra J.A., Spaan W.J.M., Zijderfeld A.J.,  
 RA Blesink-Pluym N.M.C., Hong F., van Schatzenburg G.J.M.,  
 RA Horzinek M.C., van der Zeijst B.A.M.;  
 RT "The peplimer protein sequence of the M41 strain of coronavirus IBV  
 RT and its comparison with Beaudette strains."; J. Gen. Virol. 67:2825-2831(1986).  
 CC -I- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS  
 CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M21883; AAA66575.1; -  
 DR EMBL: A24863; CA01736.1; -  
 DR PIR: S07421; S07421.  
 DR InterPro: IPR002551; Corona\_S1.  
 DR InterPro: IPR002552; Corona\_S2.  
 DR Pfam: PF01600; Corona\_S1; 1.  
 DR Pfam: PF01601; Corona\_S2; 1.  
 DR Glycoprotein; Envelope protein; Transmembrane; Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 1162  
 FT CHAIN 19 537  
 FT CHAIN 538 1162  
 FT DOMAIN 1120 1137  
 FT CARBOHYD 51 51  
 FT CARBOHYD 77 77  
 FT CARBOHYD 103 103  
 FT CARBOHYD 144 144  
 FT CARBOHYD 163 163  
 FT CARBOHYD 178 178  
 FT CARBOHYD 212 212  
 FT CARBOHYD 237 237  
 FT CARBOHYD 247 247  
 FT CARBOHYD 264 264  
 FT CARBOHYD 271 271  
 FT CARBOHYD 276 276  
 FT CARBOHYD 306 306  
 FT CARBOHYD 425 425  
 FT CARBOHYD 447 447  
 FT CARBOHYD 513 513  
 FT CARBOHYD 530 530  
 FT CARBOHYD 579 579  
 FT CARBOHYD 591 591  
 FT CARBOHYD 669 669  
 FT CARBOHYD 714 714  
 FT CARBOHYD 947 947  
 FT CARBOHYD 960 960  
 FT CARBOHYD 979 979  
 FT CARBOHYD 1014 1014  
 FT CARBOHYD 1038 1038  
 FT CARBOHYD 1051 1051  
 FT CARBOHYD 1074 1074  
 FT CARBOHYD 1162 1162  
 SQ SEQUENCE 1162 AA; 128077 MW; 3C9CCT0938492DDA CMC64;

RESULT	13	
KEMI_YEAST		
ID	STANDARD;	PRT: 1528 AA.
KEMI_YEAST		
P22147;		
01-AUG-1991	(Rel. 19, Created)	
01-AUG-1991	(Rel. 19, last sequence update)	
30-MAY-2000	(Rel. 39, last annotation update)	
STRAND EXCHANGE PROTEIN 1 (KAR-)	ENHANCING MUTATION PROTEIN) (5'-3'	
EXORIBONUCLEASE) (DNA STRAND TRANSFER PROTEIN BETA) (STP-BETA) (P175		
KEMI OR SEPI OR XANI OR DST2 OR RAR5 OR YGL173C OR G1645.		
Saccharomyces cerevisiae (Baker's yeast).		
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
Saccharomycetales; Saccharomycetaceae; Saccharomycetes;		
NCBI_TaxID=4932;		
[1]		

SEQUENCE FROM N.A.  
MEDLINE=91169260; PubMed=2076815;  
Kim J., Ijzungsahl P.O., Fink G.R.;  
"Kern mutations affect nuclear fusion in *Saccharomyces cerevisiae*." Genetics 126:799-812(1990). [2]

MEDLINE-91203880; AND PARTIAL SEQUENCE.  
 Tishkoff D., Johnson A.W., Kolodner R.D.:  
 "Molecular and genetic analysis of the gene encoding the  
 Saccharomyces cerevisiae strand exchange protein Sep1.";  
 Mol. Cell. Biol. 11:2593-2608(1991).  
 SEQUENCE FROM N.A.  
 STRAIN-A3644;  
 MEDLINE-91203879; PubMed-1850100;  
 Dykstra C.C., Kitada K., Clark A.B., Hamatake R.K., Sugano A.:  
 "Cloning and characterization of DSF2, the gene for DNA strand  
 transfer protein beta from *Saccharomyces cerevisiae*.";  
 Mol. Cell. Biol. 11:2583-2592(1991).  
 [4]

SEQUENCE FROM N.A.  
MEDLINE=91071610; PubMed=1979303;  
Larimer F.W., Stevens A.:  
"Disruption of the gene XRN1, coding for a 5'-->3' exonuclease,  
restricts yeast cell growth."; Gene 95:85-90(1990).  
[5]

SEQUENCE FROM N. A.  
MEDLINE-91227124; Pubmed-5027746;  
Kipling D., Rambini C.,  
"rar mutations which increase artificial chromosome stability in  
Saccharomyces cerevisiae identify transcription and recombination  
proteins";  
Nucleic Acids Res. 19:1385-1391(1991).  
[6]

SEQUENCE FROM N.A.  
STR001-5288C / FY1679;  
MED001-5288C / FY16931; PubMed-8619931.  
Bertalan I., Cogilevina M., Zaccaria P., Klima R., Bruschi C.V.:  
The sequence of an 11.1 kb fragment on the left arm of Saccharomyces  
cerevisiae chromosome VII reveals six open reading frames including  
NCE49, KRM1 and four putative new genes.",  
Yeast 11:1116-1194(1995).  
71

CHARACTERIZATION OF EXONUCLEASE ACTIVITY.  
MEDLINE-91310695; PubMed-1856231;  
Johnson A.W., Kolodner R.D.;  
Strand exchange protein 1 from *Saccharomyces cerevisiae*. A novel

RA  
Larimer F.W., Hsu C.L., Maupin M.K., Stevens A.;  
"Characterization of the *XN1* gene encoding a 5'->3'-  
exoribonuclease: sequence data and analysis of  
mRNA levels of gene-disrupted yeast cells.",  
Gene 120:51-57(1992).

CHARACTERIZATION OF EXONUCLEASE ACTIVITY DURING STRAND EXCHANGE.  
MEDLINE-94148871; Pubmed-8106411;  
Johnson A.W., Kolodner R.D.;  
"The activity of the Saccharomyces cerevisiae strand exchange protein  
J.1 intrinsic exonuclease during joint molecule formation.";  
J. Biol. Chem. 269:3664-3672(1994).  
[10]

111] J. Biol. Chem. 269:3573-3581(1994).  
"Characterization of the interaction of Saccharomyces cerevisiae strand exchange protein 1 with DNA.";  
Johnson A.W., Kolodner R.D.;  
MEDLINE-94148872; Pubmed-8106412;  
DNA-BINDING PROPERTIES.

CHARACTERIZATION OF EXONUCLEASE AND STRAND EXCHANGE ACTIVITIES.  
MELNINE-94245729; PubMed=8188690;  
Kaslin E., Heyer W.-D.;  
A multifunctional exonuclease from vegetative Schizosaccharomyces  
pombe cells exhibiting in vitro strand exchange activity. J.  
Biol. Chem. 269:14094-14102(1994).  
[12]

CHARACTERIZATION OF SPECIFICITY FOR G4 TETRAPLEX DNA.  
MEDLINE-94291198; Pubmed-8020096;  
Liu Z., Gilbert W.;  
"The yeast KEM1 gene encodes a nuclease specific for G4 tetraplex  
DNA: Implication of *in vivo* functions for this novel DNA structure."  
Cell 77:1083-1092(1994).  
[13]

SUBCELLULAR LOCATION  
MEDLINE=95257954; PubMed=7735953;  
Heyer W.-D., Johnson A.W., Reinhardt U., Kolodner R.D.;  
"Regulation and intracellular localization of Saccharomyces  
cerevisiae strand exchange protein 1 (Sep1/Xrn1/Kem1), a  
multifunctional exonuclease." *MOL. CELL. BIOL.* 15:12728-2736(1995).  
[14]

ASSOCIATION WITH MICROTUBULES.  
MEDLINE=95237186; PubMed=7720696;  
Interthal H., Bellocq C., Baehrel J., Baskin V.I., Edelstein S.,  
Hoyer W.-D., Bellocq C., Baehrel J., Baskin V.I., Edelstein S.,  
Sachdev R.N.S. (1995). A role of Septin (= Krim, Krim) as a microtubule-associated protein in  
EMBO J. 14:1057-1066(1995).  
FUNCTION=7720696(1995).

EXCHANGE AND 5'-3' EXONUCLEASE ACTIVITIES IN VITRO. PREFERENTIAL DEGRADES SINGLE-STRANDED DNA (SS-DNA) AND CAN REMAINE COMPLEMENTARY SS-DNA. CATALYSES THE FORMATION OF HETEROODUPLEX FROM CIRCULAR SS-DNA AND HOMOLOGOUS LINEAR DS-DNA IN VITRO. HAS BEEN SHOWN TO BE A G4-DNA-DEPENDENT EXONUCLEASE. HAS EXONUCLEASE AND A 5'->3' EXONUCLEASE.

CHROMOSOME STABILITY AND DEFECTS IN SPINDLE POLE BODY DUPLICATION AND/OR SEPARATION AS WELL AS LOSS OF VIABILITY UNDER CONDITIONS OF NITROGEN STARVATION. HOMOLOGOUS DIPLOIDS ARE UNABLE TO SPOKULATE. FUNCTION: ACTS AS A MICROTUBULE-ASSOCIATED PROTEIN WHICH INTERACTS WITH CYTOSOLIC MICROTUBULES THROUGH BETA-TUBULIN. MAGNESIUM: BOTH STRAND EXCHANGE AND NITROGEN

1- SUBCELLULAR LOCATION: PREDOMINANTLY CYTOPLASMIC, ALSO PERINUCLEAR

-1- SIMILARITY: STRONG TO S.POMBE EXO2; SOME, TO YEAST RAT1 AND TO S.POMBE DHPI.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL: M90097; AAA35219.1; -  
DR EMBL: X54717; CAA38520.1; -  
DR EMBL: M58367; AAA55036.1; -  
DR EMBL: M36725; AAA5125.1; -  
DR EMBL: X61181; CAA43487.1; -  
DR EMBL: X84705; CAA59180.1; -  
DR EMBL: Z72695; CAA96885.1; -  
DR PIR: S13743; S13743.  
DR PIR: S16885; S16885.  
DR PIR: S16701; S16701.  
DR PIR: A39790; A39790.  
DR SGP: S0003141; KEM1.  
KW Hydrolyase; Nuclease; Exonuclease; Magnesium; DNA recombination;  
KW DNA-binding; DNA damage.  
SQ SEQUENCE 1528 AA; 175459 MW; 49C2EDAF73D3EB92 CRC64;

Query Match 42.7%; Score 41; DB 1; Length 1528;  
Best Local Similarity 60.0%; Pred. No. 67;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 HCNFNDVTRLENE 17  
| : ||| ||| |  
DB 45 HCNDDVTKRLTEEE 59

RESULT 14  
RPO\_CGMVS STANDARD: PRT: 1648 AA.  
ID RPO\_CGMVS STANDARD: PRT: 1648 AA.  
AC P19523; O83208; P89877; P90356.  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (186 KDA PROTEIN) [CONTAINS: METHYLTRANSFERASE/RNA HELICASE (MT/HEL) (129 KDA PROTEIN)].  
OS Cucumbers green mottle mosaic virus (watermelon strain SH) (CGMV), and  
OS Cucumbers green mottle mosaic virus (watermelon strain W) (CGMW).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.  
OX NCBI\_TaxID=12236; 12237;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN-SH:  
RC MEDLINE=91311400; PubMed=1856687;  
RA Ugaiki M., Tomiyama M., Kakutani T., Hidaka S., Kiguchi T., Nagata R., Sato T., Motoyoshi F., Nishiguchi M.;  
RT "The complete nucleotide sequence of cucumber green mottle mosaic virus (SH strain) genomic RNA."  
RL J. Gen. Virol. 72:1487-1495(1991).  
[2]  
RN REVISIONS.  
RC STRAIN-SH:  
RA Ugaiki M., Tomiyama M., Kakutani T., Hidaka S., Kiguchi T., Nagata R., Sato T., Motoyoshi F., Nishiguchi M.;  
RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.  
[3]  
RN SEQUENCE OF 1496-1648 FROM N.A.  
RP STRAIN-W;  
RC MEDLINE=89073773; PubMed=3201760;  
RA Sato T., Imai Y., Meshi T., Okada Y.;  
RT "Intertrial homologues of the 30k proteins of tobamoviruses."  
RL Virology 167:653-656(1988).  
-1- FUNCTION: THE LONGER PROTEIN IS A RNA POLYMERASE ACTIVE IN VIRAL

RNA REPLICATION.

-1- FUNCTION: THE SMALLER PROTEIN IS A METHYLTRANSFERASE ACTIVE IN RNA CAPPING AND AN RNA HELICASE.

CC MISCELLANEOUS: READTHROUGH OF THE TERMINATOR UAG OCCURS BETWEEN COODONS FOR LYS-1144 AND GLN-1146.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL: D12505; BAA18895.1; -  
DR EMBL: D12505; BAA18896.1; -  
DR EMBL: J04322; AAA46382.1; -  
DR PIR: JQ1157; WMTMS2.  
DR InterPro: IPR001788; RNA\_dep\_RNAPol2.  
DR InterPro: IPR002588; V\_methyltransf.  
DR InterPro: IPR006006; Viral\_helcsl.  
DR Pfam: PF00978; RNA\_dep\_RNAPol2; 1.  
DR Pfam: PF01443; Viral\_helicase1; 1.  
DR Pfam: PF01660; Vmethyltransf; 1.  
KW Transferase; RNA-directed RNA polymerase; Helicase; ATP-binding.  
FT CHAIN 1 1648  
FT CHAIN 1 1144  
KW RNA-DIRECTED RNA POLYMERASE.  
FT NP\_BIND 863 870  
FT NP\_BIND 863 870  
SQ SEQUENCE 1648 AA; 186549 MW; 1D1AFEEEB7B65595 CRC64;

Query Match 42.7%; Score 41; DB 1; Length 1648;  
Best Local Similarity 61.5%; Pred. No. 72;  
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 5 NENDVTRLENE 17  
: | : ||| ||| :  
DB 1188 DEFDVTRLRDNE 1200

RESULT 15  
UTRO\_HUMAN STANDARD: PRT: 3433 AA.  
ID UTRO\_HUMAN STANDARD: PRT: 3433 AA.  
AC P46939;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE UTRORPHIN (DYSTROPHIN-RELATED PROTEIN 1) (DRP1) (DRP).  
GN UTRN OR DMDL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=93096045; PubMed=1461283;  
RA Tinsley J.M., Blake D.J., Roche A., Fairbrother U., Riss J., Byth B.C., Knight A.E., Kendrick-Jones J., Suthers G.K., Love D.R., Edwards Y.H., Davies K.E.;  
RT "Primary structure of dystrophin-related protein."  
RL Nature 360:591-593(1992).  
[2]  
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 144-261.  
RP MEDLINE=99141377; PubMed=9887274;  
RA Keep N.H., Norwood F.L.M., Moores C.A., Winder S.J., Kendrick-Jones J.;  
RT "The 2.0-A structure of the second calponin homology domain from the actin-binding region of the dystrophin homologue utrophin."  
RL J. Mol. Biol. 285:1257-1264(1999).  
-1- FUNCTION: MAY PLAY A ROLE IN ANCHORING THE CYTOSKELETON TO THE PLASMA MEMBRANE (BY SIMILARITY TO DYSTROPHIN).  
-1- SUBCELLULAR LOCATION: NEURONMUSCULAR JUNCTION.  
-1- TISSUE SPECIFICITY: MUSCLE.

CC -1- SIMILARITY: STRONG, TO DYSTROPHIN.  
 CC -1- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY  
 CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN,  
 CC ABP-120, ABP-180, OR BETA-FODRIN).  
 CC -1- SIMILARITY: CONTAINS 1 WW DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 20 SPECTRIN REPEATS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X69086; CAA48829.1; -  
 DR PIR: S28381; S28381.  
 DR PDB: 1BHD; 16-FEB-99.  
 DR MIM: 128240; -  
 DR InterPro: IPR001589; Actinin\_act\_bind.  
 DR InterPro: IPR001715; Calponin\_hom.  
 DR InterPro: IPR002017; Spectrin.  
 DR InterPro: IPR003122; Tarrh.  
 DR InterPro: IPR001202; WW.  
 DR InterPro: IPR002349; WW\_Domain.  
 DR Pfam: PF00307; CH; 2.  
 DR Pfam: PF00307; spectrin; 19.  
 DR Pfam: PF00337; WW; 1.  
 DR Pfam: PF00369; Z2; 1.  
 DR PRINTS: PR00403; WMDOMAIN.  
 DR SMART: SM00033; CH; 2.  
 DR SMART: SM00150; SPEC; 18.  
 DR SMART: SM00319; Tarrh; 1.  
 DR SMART: SM00456; WW; 1.  
 DR PROSITE: PS00019; ACTININ\_1; 1.  
 DR PROSITE: PS00020; ACTININ\_2; 1.  
 DR PROSITE: PS00021; CH; 2.  
 DR PROSITE: PS01159; WW\_DOMAIN\_1; 1.  
 DR PROSITE: PS00020; WW\_DOMAIN\_2; 1.  
 DR KW structural protein; Actin-binding; Calcium-binding; Cytoskeleton;  
 KW Repeat; 3D-structure.  
 FT DOMAIN 1 246  
 FT REPEAT 246 364 ACTIN-BINDING.  
 FT REPEAT 309 417 SPECTRIN 1.  
 FT REPEAT 418 526 SPECTRIN 2.  
 FT REPEAT 541 637 SPECTRIN 3.  
 FT REPEAT 687 798 SPECTRIN 4.  
 FT REPEAT 803 902 SPECTRIN 5.  
 FT REPEAT 1016 1083 SPECTRIN 6.  
 FT REPEAT 1125 1230 SPECTRIN 7.  
 FT REPEAT 1248 1334 SPECTRIN 8.  
 FT REPEAT 1432 1541 SPECTRIN 9.  
 FT REPEAT 1544 1649 SPECTRIN 10.  
 FT REPEAT 1652 1753 SPECTRIN 11.  
 FT REPEAT 1910 1968 SPECTRIN 12.  
 FT REPEAT 1976 2081 SPECTRIN 13.  
 FT REPEAT 2258 2333 SPECTRIN 14.  
 FT REPEAT 2399 2440 SPECTRIN 15.  
 FT REPEAT 2443 2556 SPECTRIN 16.  
 FT REPEAT 2559 2636 SPECTRIN 17.  
 FT REPEAT 2658 2688 SPECTRIN 18.  
 FT REPEAT 2691 2797 SPECTRIN 19.  
 FT DOMAIN 2812 2845 SPECTRIN 20.  
 FT DOMAIN 3064 3111 WW.  
 FT DOMAIN 3117 3177 Z2.  
 FT SEQUENCE 3433 AA; 394488 MW; EA88BD409F8585B CRC64;  
 CYS-RICH.

Query Match  
 Best Local Similarity 42.7%; Score 41; DB 1; Length 3433;  
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 HCNENDVYTRLENE 17  
 Db 1982 HCDLMDLTOWITEAE 1996  
 ||: ||: | | |  
 RESULT 16  
 ID R1MM\_RICPR STANDARD; PRT; 165 AA.  
 AC Q9ZD10;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DE PROBABLE 16S RNA PROCESSING PROTEIN R1MM.  
 GN R1MM OR RP348.  
 OS Rickettsia prowazekii.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
 OC Rickettsiaceae; Rickettsiae; Rickettsia.  
 OX NCBI\_TaxID=782;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MADRID E;  
 RA MEDLINE-99039499; PubMed-9823893;  
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,  
 RA Sierheltz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,  
 RA Eriksson A.-S., Winkler H.H., Kuitland C.G.;  
 RA "The genome sequence of Rickettsia prowazekii and the origin of  
 RA mitochondria.";  
 RL Nature 396:133-140(1998).  
 CC -1- FUNCTION: ESSENTIAL FOR EFFICIENT PROCESSING OF 16S RNA. PROBABLY  
 CC PART OF THE 30S SUBUNIT PRIOR TO OR DURING THE FINAL STEP IN THE  
 CC PROCESSING OF 16S FREE 30S RIBOSOMAL SUBUNITS. IT COULD BE SOME  
 CC ACCESSORY PROTEIN NEEDED FOR EFFICIENT ASSEMBLY OF THE 30S  
 CC SUBUNIT. R1MM IS NEEDED IN A STEP PRIOR TO RBEA DURING THE  
 CC MATURATION OF 16S RNA. HAS AFFINITY FOR FREE RIBOSOMAL 30S  
 CC SUBUNITS BUT NOT FOR 70S RIBOSOMES (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE R1MM FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AJ235271; CAA14808.1; -  
 DR InterPro: IPR002676; R1MM.  
 DR Pfam: PF01782; R1MM; 1.  
 DR KW RNA processing; Complete proteome.  
 DR RNA processing; Complete proteome.  
 DR SEQUENCE 165 AA; 18772 MW; 25DDB377BC250919 CRC64;

Query Match  
 Best Local Similarity 41.7%; Score 40; DB 1; Length 165;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CNEFDVYTR 12  
 Db 63 CTENDIATR 71  
 ||: ||: | | |  
 RESULT 17  
 ID HFA3\_HAEIN STANDARD; PRT; 212 AA.  
 AC P43988;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DE MAJOR FIMBRIAL SUBUNIT PRECURSOR (PILIN).  
 GN HIFA.  
 OS Haemophilus influenzae.

CC	Bacteroidetes; Proteobacteria; gamma subdivision; Pasteurellales;
OC	Hemophilus.
OX	NCBI_TaxID=727;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-86-0295 / LKP SEROTYPE 1;
RA	Green B.A., Olmsted S.B.;
RL	Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC	-1- FUNCTION: MEDIATES ADHERENCE TO OROPHARYNGEAL EPITHELIAL CELLS.
CC	HELPS THE AIRWAY COLONIZATION PROCESS.
CC	-1- SIMILARITY: BELONGS TO THE FIMBRIA/PAPA FAMILY OF FIMBRIA PROTEINS.
CC	-1- SIMILARITY: BELONGS TO THE FIMBRIA/PAPA FAMILY OF FIMBRIA PROTEINS.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> )
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> .
DR	EMBL; U19730; AAA61814.1; -
DR	InterPro; IPR000259; Fimbrin1.
DR	Pfam; PF00419; Fimbrin1; 1.
KW	Fimbrin; Signal.
FT	SIGNAL 1
FT	CHAIN 18
FT	DISULFID 42
SE	SEQUENCE 212 AA; 23003 MW; 3FAAB87B46901B1 CRC64;
	POTENTIAL MAJOR FIMBRIN SUBUNIT. PROBABLE.

Query Match	41.7%	Score 40;	DB 1;	Length 212;
Best Local Similarity	47.1%	Pred. No. 12;		
Matches	8;	Conservative	2;	Mismatches 7;
				Indels 0;
				Gaps 0;
OY	1	FEHCNFNDVTRLRENE	17	
Db	151	FTTHSTNDVATQQTIVNK	167	

```

RESULT 18
FLPA_SULAC
ID FLPA_SULAC STANDARD: PRT: 231 AA.
AC QP99M0:
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE FIBRILLARIN-LIKE PRE-RRNA PROCESSING PROTEIN.
GN FLPA OR TIB.
OS Sulfolobus acidocaldarius.
OC Archaeae: Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2285;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20237948; PubMed=10775111;
RA Omer A.D., Lowe T.M., Russell A.G., Ehardt H., Eddy S.R.,
RA Dennis P.P.;
RL "Homologs of small nucleolar RNAs in Archaea.";
RL FUNCTION: MAY BE INVOLVED IN PRE-RRNA PROCESSING.
CC -1- SIMILARITY: BELONGS TO THE FIBRILLARIN FAMILY.
CC -1- SIMILARITY: BELONGS TO THE FIBRILLARIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AF201093; AAE69254.1; -
CC InterPro: IPR000692; FibrillarIn.
CC Pfam: PF01269; FibrillarIn.1.
CC -----

```

DR PRINTS; PR00052; FIBRILLARIN.  
DR PRODOM; PD004637; Fibrillarin; 1.  
DR PROSITE; PS00566; FIBRILLARIN; 1.  
KW rRNA processing; RNA-binding.  
SQ SEQUENCE 231 AA; 26445 MW; 430DC184CB2EC7B CRC64;

Query Match	41.7%	Score 40;	DB 1;	Length 231;
Best Local Similarity	80.0%	Pred. No. 13;		
Matches	8;	Conservative	0;	Mismatches
			2;	Indels
				Gaps
				0;
Qy	4	CNFNDVTRL	13	
Db	18	CIENGDTRL	27	

RESULT	19
KMS1_SCHPO	
ID	KMS1_SCHPO
AC	p87245
DT	20-AUG-2001 (Rel. 40, Created)
DT	20-AUG-2001 (Rel. 40, Last sequence update)
DT	20-AUG-2001 (Rel. 40, Last annotation update)
DE	KARYOGAMY MEIOTIC SEGREGATION PROTEIN 1.
GN	KMS1 OR SPAC3A11.05C.
OS	Schizosaccharomyces pombe (Fission yeast).
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;
OC	Schizosaccharomyces.
OX	NCBI_TaxID:4896;

RP SEQUENCE FROM N.A. AND FUNCTION  
RX MEDLINE=97294501; PubMed=9150257;  
RA Shimanuki M., Miki F., Ding D.-Q., Chikashige Y., Hiraoka Y.,  
RA Horio T., Nima O.,  
RT "A novel fission yeast gene, *kms1+*, is required for the formation of  
RT meiotic prophase-specific nuclear architecture."  
RL Mol. Gen. Genet. 254:238-249(1997).

RC STRAIN-972;  
RA Oliver K., Harris D., Barrell B. G., Rajadream M. A., Wood V.;  
RL submitted (SEP-1997) to the EMBL/GenBank/DBD databases.  
CC - FUNCTION: HAS A ROLE IN KAROGAMY, RECOMBINATION AND SEGREGATION  
CC DURING MEIOSIS. ALTHOUGH IT HAS BEEN SHOWN TO ASSOCIATE WITH THE  
CC SPINDLE POLE BODY IT IS UNLIKELY TO BE INVOLVED IN ITS FORMATION  
CC OR MAINTENANCE SINCE KMS1(-) MUTANTS ARE ABLE TO COMPLETE MITOSIS  
CC SUCCESSFULLY.

CC This SWISSPROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; D84439; BAA20460.1; -  
DR EMBL; 299260; CAB16381.1; -  
KW Cell division; Meiosis;  
SQ SEQUENCE 607 AA; 69240 MW; 216A1D5CA93C9550 CRC64

Query Match	41.7%	Score 40	DB 1	Length 607
Best Local Similarity	61.58%	Pred. No. 37		
Matches 8; Conservative		1; Mismatches	4; Indels	0; Gaps

```
QY      5 NENDVTRLRENE 17
         ||| : |||
Db      70 NFRDVKKLTENE 82
```

## RESULT 20

HS83\_DROAV  
ID HS83\_DROAV STANDARD: PRT: 716 AA.  
AC 002192;  
DT 15-DEC-1998 (Rel. 37, Created)  
DE 15-DEC-1998 (Rel. 37, Last sequence update)  
HEAT SHOCK PROTEIN 83 (HSP 82).  
GN HSP83.  
OS Drosophila auraria (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
RN NCBI\_TaxID=47315;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3040.11B;  
RA Konstantopoulos I., Scouras Z.G.;  
RT "The hsp83 gene of Drosophila auraria";  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: MOLECULAR CHAPERONE. HAS ATPASE ACTIVITY.  
CC (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL: U75687; AAB58358.1;  
DR FlyBase; FBgn0020208; Daur\Hsp83.  
DR InterPro; IPR003594; HATPase\_C.  
DR Pfam; PF02518; HATPase\_C; 1.  
DR Pfam; PF00183; HSP90; 1.  
DR PRINTS; PR00775; HEATSHOCK90.  
DR SMART; SM00387; HATPase\_C; 1.  
DR PROSITE; PS00298; HSP90; 1.  
DR Chaperone; ATP-binding; Heat shock.  
SQ SEQUENCE 716 AA; 81760 MW; 9EC59F027C4DFCD7 CRC64;  
  
Query Match  
Best Local Similarity 41.7%; Score 40; DB 1; Length 716;  
Matches 5; Conservative 7; Mismatches 4; Indels 0; Gaps 0;  
  
QY 2 EHCNENDVTTRENE 17  
Db 457 DFCSLSDYVRMKENQ 472

RESULT 21  
SV2\_RAT  
ID SV2\_RAT STANDARD: PRT: 742 AA.  
AC 002563;  
DT 01-FEB-1995 (Rel. 31, Created)  
DE 30-MAY-2000 (Rel. 39, Last sequence update)  
SYNAPTIC VESICLE PROTEIN 2 (SV2).  
GN SV2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
RN NCBI\_TaxID=10116;  
[1]  
RP SOURCE FROM N.A., AND SEQUENCE OF 1-40.  
RC TISSUE=Brain;  
RX MEDLINE=92390722; PubMed=1519064;  
RA Bajajalleh S.M., Peterson K., Shingal R., Scheller R.H.;  
RT "SV2, a brain synaptic vesicle protein homologous to bacterial

RT transporters";  
RL Science 257:1271-1273(1992).  
CC -1- FUNCTION: MAY ACT AS A VESICLE-LOCALIZED NEUROTRANSMITTER  
CC TRANSPORTER OR MAY FUNCTION AS AN ION TRANSPORTER OR CHANNEL.  
CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLE.  
CC -1- TISSUE SPECIFICITY: NEURAL AND ENDOCRINE CELLS OF BRAIN AND  
CC SPINAL CORD.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL: L05435; AAA42188.1;  
DR Pfam; PF00083; sugar-1;  
KW Synapse; Nerve; Glycoprotein; Neurotransmitter transport;  
KW Transmembrane.  
FT DOMAIN 1  
FT TRANSMEM 164 163  
FT DOMAIN 183 182  
FT TRANSMEM 205 204  
FT DOMAIN 225 225  
FT TRANSMEM 243 242  
FT DOMAIN 261 262  
FT TRANSMEM 263 281  
FT DOMAIN 282 284  
FT TRANSMEM 295 322  
FT DOMAIN 323 334  
FT TRANSMEM 335 355  
FT DOMAIN 356 445  
FT TRANSMEM 446 469  
FT DOMAIN 470 592  
FT TRANSMEM 593 611  
FT DOMAIN 612 626  
FT TRANSMEM 627 647  
FT DOMAIN 648 649  
FT TRANSMEM 650 669  
FT DOMAIN 670 694  
FT TRANSMEM 695 711  
FT DOMAIN 712 712  
FT TRANSMEM 713 731  
FT DOMAIN 732 742  
FT CARBOHYD 498 498  
FT CARBOHYD 548 548  
FT CARBOHYD 573 573  
SQ SEQUENCE 742 AA; 82705 MW; 565DE7EF2929D5DB CRC64;  
  
Query Match  
Best Local Similarity 41.7%; Score 40; DB 1; Length 742;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 FEHCNENDVTT 11  
Db 531 FEECYFEDVTS 541

RESULT 22  
Y594\_METUA  
ID Y594\_METUA STANDARD: PRT: 169 AA.  
AC 058012;  
DT 01-NOV-1997 (Rel. 35, Created)  
DE 20-NOV-1997 (Rel. 35, Last sequence update)  
DE HYPOTHEICAL PROTEIN M00594.  
GN M00594.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
OC Methanococcus.  
RN NCBI\_TaxID=2190;  
[1]

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE-9633799; PubMed-8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Overbeek A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,  
 RA Kierstead R., Kirkness E.F., Weissbrock K.G., Merrick J.M., Glöckle A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,  
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*  
 RT *jannaschii*.";  
 RL Science 273:1058-1073(1996).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: 067508; AAB9595.1; -  
 CC TIGR: M0594; -  
 DR InterPro: IPR002799; IMP4.  
 DR Pfam: PF01945; IMP4; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 169 AA; 20053 MW; 9FF6090865C6E4B0 CRC64;

Query Match 41.1%; Score 39.5; DB 1; Length 169;  
 Best Local Similarity 42.1%; Pred. No. 11;  
 Matches 8; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

QY 1 FEHCNFN--DVTTRLREN 16  
 | | | | | | | | | |  
 Db 119 FOHLNINEDSITRLREKD 137

RESULT 23  
 ID SPOT\_SYNY3 STANDARD; PRT; 760 AA.  
 AC P74007;  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE PROBABLE GUANOSINE-3',5'-BIS(DIPHOSPHATE) 3'-PYROPHOSPHOHYDROLASE  
 DE (EC 3.1.7.2) ((PGPP)ASE) (PENTA-PHOSPHATE GUANOSINE-3'-  
 DE PYROPHOSPHOHYDROLASE).  
 GN SPOT OR STR1325.  
 OS *Synechocystis* sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; *Synechocystis*.  
 OX NCBI\_TaxID=1148;  
 RN [1]  
 RX MEDLINE-97061201; PubMed-8905231;  
 RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyajima N., Hirotsawa M., Sugita M., Sasamoto S., Kimura T.,  
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
 RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M.,  
 RA Tabata S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT *Synechocystis* sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";  
 RL DNA Res. 3:109-136(1996).  
 CC -1- FUNCTION: IN EUKARYOTIC PGPP (GUANOSINE 3'-DIPHOSPHATE 5'-  
 CC DIPHOSPHATE) IS A MEDIATOR OF THE STRINGENT RESPONSE THAT  
 CC COORDINATES A VARIETY OF CELLULAR ACTIVITIES IN RESPONSE TO  
 CC CHANGES IN NUTRITIONAL ABUNDANCE. THIS ENZYME CATALYZES THE  
 CC DEGRADATION OF PGPP INTO GDP. IT MAY ALSO BE CAPABLE OF  
 CC CATALYZING THE SYNTHESIS OF PGPP (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: GUANOSINE-3',5'-BIS(DIPHOSPHATE) + H(2)O =  
 CC GUANOSINE-5'-DIPHOSPHATE + PYROPHOSPHATE.  
 CC -1- COFACTOR: MANGANESE (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE RELA / SPOT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D90911; BAA18078.1; -  
 DR InterPro: IPR002912; ACT.  
 DR InterPro: IPR003607; HDC.  
 DR Pfam: PF01842; ACT; 1.  
 DR SMART: SM00471; HDC; 1.  
 KW Hydrolase; Manganese; Complete proteome.  
 SQ SEQUENCE 760 AA; 86568 MW; CA276EA2286358F7 CRC64;

Query Match 41.1%; Score 39.5; DB 1; Length 760;  
 Best Local Similarity 31.2%; Pred. No. 57;  
 Matches 10; Conservative 1; Mismatches 4; Indels 17; Gaps 1;

QY 2 EHCFN-----NDVTTRLREN 16  
 | | | | | | | | | |  
 Db 548 EHCYQVVEDLACLGYEITSNSVNRKRN 579

RESULT 24  
 ID ENV\_STVYZ STANDARD; PRT; 854 AA.  
 AC P17281;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE ENVELOPE POLYPROTEIN GP160 PRECURSOR [CONTAINS: EXTERIOR MEMBRANE  
 DE GLYCOPROTEIN (GP120); TRANSMEMBRANE GLYCOPROTEIN (GP41)].  
 GN ENV.  
 OS Chimpanzee immunodeficiency virus (STV/cpz) (CIV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11723;  
 RN [1]  
 RX MEDLINE-90259077; PubMed-2188136;  
 RA Hueb T., Cheynier R., Meyers A., Roelants G., Wain-Hobson S.;  
 RT "Genetic organization of a chimpanzee lentivirus related to HIV-1.";  
 RL Nature 345:356-359(1990).  
 CC -1- SIMILARITY: STRONG, WITH HIV-1 ENV POLYPROTEIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X52154; CAA36407.1; -  
 CC PIR: S09990; VCLJST.  
 DR HIV: X52154; ENVSCP2.  
 DR InterPro: IPR000328; ENV GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polypeptide; Polypeptide; Transmembrane;  
 KW Signal.  
 FT SIGNAL 1 30  
 FT CHAIN 31 500 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 501 854 TRANSMEMBRANE GLYCOPROTEIN.  
 FT TRANSMEM 501 517 POTENTIAL.

```

FT TRANSMEM 675 693 POTENTIAL.
FT TRANSMEM 805 821 POTENTIAL.
FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 426 426 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 854 AA; 95803 MW; 2E249AFAD4F3D9B3 CRC64;

```

```

Query Match 41.1%; Score 39.5; DB 1; Length 854;
Best Local Similarity 50.0%; Pred. No. 64;
Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

```

```

Db 153 KNCSEF-VTTLEJDKK 167
2 EHCNPNVYTRIRENE 17
:::|||||:::
ID CD59_PIG STANDARD; PRT; 123 AA.
AC 062680: 09X94; 09TR76;
DT 15-DEC-1998 (Rel. 37, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE CD59 GLYCOPROTEIN PRECURSOR (MEMBRANE ATTACK COMPLEX INHIBITION
  FACTOR) (MACIF) (MAC-INHIBITORY PROTEIN) (MAC-IP) (PROTECTIN).
GN CD59.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN NCBI_TaxID:9823;
RP [1]
RC SEQUENCE FROM N.A.
RA MEDLINE:98217182; PubMed:9558099;
RT Hinchliffe S.J.; Rushmere N.K.; Hanna S.M.; Morgan B.P.;
  "Molecular cloning and functional characterization of the pig
  analogue of CD59: relevance to xenotransplantation.";
  J Immunol. 160:3924-3932(1998).
RP [2]
RC SEQUENCE OF 26-123 FROM N.A., AND FUNCTION.
RA TISSUE-Aortic endothelium;
RX MEDLINE:99023683; PubMed:9808497;
RA Maher S.E.; Pflugh D.L.; Larsen N.J.; Rothschild M.F.;
  "Structure/function characterization of porcine CD59: expression,
  RT chromosomal mapping, complement-inhibition, and costimulatory
  activity.";
  Transplantation 66:1094-1100(1998).
RN [3]

```

```

RP SEQUENCE OF 26-64.
RC TISSUE-Erythrocyte;
RX MEDLINE:95181826; PubMed:753195;
RA van den Berg C.W.; Harrison R.A.; Morgan B.P.;
  "A rapid method for the isolation of analogues of human CD59 by
  RT preparative SDS-PAGE: application to pig CD59.";
  J. Immunol. Methods 179:223-231(1995).
RL j. Immunol. Methods 179:223-231(1995).
CC -1- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK
  COMPLEX (MAC) ACTION. ACTS BY BINDING TO THE C8 AND/OR C9
  CC COMPLEMENTS OF THE ASSEMBLING MAC, THEREBY PREVENTING
  CC INCORPORATION OF THE MULTIPLE COPIES OF C9 REQUIRED FOR COMPLETE
  CC FORMATION OF THE OSMOTIC PORE.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
  CC LIVER, KIDNEY, SPLEEN, HEART AND SKELETAL MUSCLE. HIGHEST LEVELS
  CC IN LUNG AND SPLEEN, LOWEST LEVELS IN LIVER AND SKELETAL MUSCLE.
CC -1- SIMILARITY: CONTAINS 1 UPAR/LY6 DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
  CC use by non-profit institutions. There are no restrictions on its
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF020302; AAC67231.1; -
CC DR InterPro: IPR001526; LY6-UPAR.
CC DR Pfam: PF00021; UPAR_LY6; 1.
CC DR SMART: SM00134; LY6; 1.
CC DR PROSITE: PS00983; LY6-UPAR; 1.
CC KW Antigen; Glycoprotein; GPI-anchor; Signal.
CC FT CHAIN 1 25
CC FT PROPEP 26 98
CC FT DOMAIN 26 103
CC FT DISULFID 28 51
CC FT DISULFID 31 38
CC FT DISULFID 44 65
CC FT DISULFID 71 89
CC FT CARBOHYD 90 95
CC FT LIPID 43 43
CC FT CONFLICT 26 98
CC FT CONFLICT 46 46
CC FT CONFLICT 63 63
CC SQ SEQUENCE 123 AA; 13790 MW; FDE2ED3F93C91321 CRC64;

```

```

Search completed: April 2, 2002, 09:23:23
Job time: 187 sec

```





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: April 2, 2002, 09:20:36 ; Search time 22.85 Seconds

(without alignments)  
108.824 Million cell updates/sec

Title: US-09-020-393B-3\_COPY\_42\_58  
Perfect score: 96  
Sequence: 1 FEHCNFDVTRLENE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 50 summaries

## Database :

SPTREMBL\_17:\*  
1: sp-archaea:\*  
2: sp-bacteria:\*  
3: sp-fungi:\*  
4: sp-human:\*  
5: sp-invertebrate:\*  
6: sp-mammal:\*  
7: sp-mhc:\*  
8: sp-organelle:\*  
9: sp-phage:\*  
10: sp-plant:\*  
11: sp-protent:\*  
12: sp-virus:\*  
13: sp-vertebrate:\*  
14: sp-unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	50.0	721	5	Q27308 anopheles a
2	46.5	48.4	1767	5	Q19346 caenorhabd
3	46	47.9	284	2	Q915M0 pseudomonas
4	46	47.9	777	2	Q9KER0 bacillus ba
5	45	46.9	1307	5	Q93554 caenorhabd
6	44	45.8	205	2	Q9ZCG3 rickettsia
7	44	45.8	513	12	Q82618 avian infec
8	44	45.8	530	12	O55343 avian infec
9	44	45.8	807	1	O27791 methanobact
10	44	45.8	1101	5	O9WD03 ciona intes
11	43	44.8	544	12	O82667 avian infec
12	43	44.8	545	12	O82619 avian infec
13	43	44.8	739	5	O9V9E6 dirosophila
14	42.5	44.3	96	12	Q70950 human immun
15	42.5	44.3	240	12	P90081 human immun
16	42.5	44.3	240	12	P90082 human immun
17	42.5	44.3	240	12	P90085 human immun
18	42.5	44.3	240	12	P90080 human immun
19	42.5	44.3	240	12	P90083 human immun

20	42.5	44.3	313	2	Q9S091 borrelia bu
21	42.5	44.3	679	12	Q9IK02 human immun
22	42.5	44.3	684	12	Q9IK06 human immun
23	42.5	44.3	852	12	Q73350 human immun
24	42.5	44.3	861	12	Q9IK05 human immun
25	42.5	44.3	861	12	Q9IK04 human immun
26	42.5	44.3	861	12	Q9IK03 human immun
27	42.5	44.3	862	12	Q9YK01 human immun
28	42	43.8	290	8	O9G157 humulic acid
29	42	43.8	302	2	O9LBM1 clostridium
30	42	43.8	332	8	O9TK28 pictetia ac
31	42	43.8	339	8	O9TK29 ormocarpum
32	42	43.8	365	5	O9XTE4 caenorhabd
33	42	43.8	395	8	O9TK00 ormocarpum
34	42	43.8	419	2	O9KG62 bacillus ba
35	42	43.8	514	8	O9TK31 ormocarpops
36	42	43.8	514	8	O9TK30 ormocarpops
37	42	43.8	514	8	O9TK29 ormocarpum
38	42	43.8	514	8	O9TK28 ormocarpum
39	42	43.8	514	8	O9TK27 ormocarpum
40	42	43.8	514	8	O9TK26 ormocarpum
41	42	43.8	514	8	O9TK25 ormocarpum
42	42	43.8	514	8	O9TK23 diphyssa flo
43	42	43.8	514	8	O9TK19 pictetia an
44	42	43.8	514	8	O9TK01 diphyssa an
45	42	43.8	573	5	O16166 dirosophila
46	41.5	43.2	87	12	O70912 human immun
47	41.5	43.2	95	12	O9YK30 human immun
48	41.5	43.2	328	12	O9IU90 human immun
49	41.5	43.2	332	12	O9IU01 human immun
50	41.5	43.2	356	12	P87970 human immun

## ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	721 AA.
Q27308				
AC	Q27308:			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	HEAT SHOCK PROTEIN 82.			
GN	HSP82.			
OS	Anopheles albimanus (New world malaria mosquito).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;			
OC	Anopheles.			
OX	NCBI_TaxID=7167;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SANTRA TECLA;			
RX	MEDLINE=96209490; PubMed=8630537;			
RA	Benedict M.O., Levine B.J., Ke Z.X., Cockburn A.F., Seawright J.A.;			
RT	'Precise limitation of concerted evolution to ORFs in mosquito Hsp82			
RT	genes. Mol. Biol. 5:73-79(1996).			
RL	Insect Mol. Biol. 5:73-79(1996).			
EMBL	LA7285; AAB05638.1; -.			
DR	EMBL; LA7285; AAB05638.1; -.			
DR	HSSP; P07900; IYER.			
DR	InterPro; IPR003594; HATPase_C.			
DR	InterPro; IPR001404; HSP90.			
DR	Pfam; PF002518; HATPase_C; 1.			
DR	Pfam; PF00183; HSP90; 1.			
DR	PRINTS; PR00775; HEATSHOCK90.			
DR	SMART; SM00387; HATPase_C; 1.			
DR	PROSITE; PS00298; HSP90; 1.			
KW	Heat shock.			
SO	SEQUENCE 721 AA; 82153 MW; C71867C5610452EA CRC64;			

Query Match 50.0%; Score 48; DB 5; Length 721;

Best Local Similarity 43.8%; Pred. No. 9.1;

Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 EHCNPNVYTRRLRENE 17  
Db 455 EYCSLNDYVGRKRENO 470

## RESULT 2

ID 019346 PRELIMINARY; PRT; 1767 AA.  
AC 019346;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
GN F1IC1.5 PROTEIN.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Palmer S.;  
RL Submitted (SEP-1995) to the EMBL/Genbank/DBJ databases.  
DR EMBL: Z54270; CA91030.1; -  
DR EMBL: Z54270; CA91031.1; -  
DR InterPro: IPR000517; Ribosomal\_L30.  
DR InterPro: IPR002035; VFMA.  
DR PROSITE: PS00634; RIBOSOMAL\_L30; UNKNOWN\_1.  
DR SMART: SM00324; VFMA; 1.  
KW Alternative splicing.  
FT VARSPPLIC 1301 1302  
FT VARSPPLIC 1303 1767  
SQ SEQUENCE 1767 AA; 200015 MW; A113DCB86F5FE85 CRC64;

Query Match  
Best Local Similarity 54.5%; Score 46.5; DB 5; Length 1767;  
Matches 12; Conservative 1; Mismatches 4; Indels 5; Gaps 2;

QY 1 FEH----CN-FNDVYTRRLRENE 17  
Db 601 FEHHLNANVFNDDTIRTKRENE 622

## RESULT 3

ID 0915M0 PRELIMINARY; PRT; 284 AA.  
AC 0915M0;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE PROBABLE TRANSCRIPTIONAL REGULATOR.  
GN PA0708.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subphylum; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Hickey C.K., Pham X.-O.T., Erwin A.L., Mitochnich S.D., Watrenner P.,  
Barber R.L., Goltzer L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
Feiler J., Selzer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT Complete genome sequence of Pseudomonas aeruginosa PA01, an  
opportunistic pathogen.";  
RL Nature 406:959-964(2000);  
RN

CC -1 SIMILARITY: TO THE LYSR FAMILY OF TRANSCRIPTIONAL REGULATORS.  
DR EMBL: AE004506; AAG04097.1; -  
DR InterPro: IPR000847; HTH\_LysR.  
DR Pfam: PF00126; HTH\_1; 1.  
DR PROSITE: PS00044; HTH\_LYSR\_FAMILY; UNKNOWN\_1.  
KW Complete proteome; DNA-binding; Transcription regulation.  
SQ SEQUENCE 284 AA; 31527 MW; 1F986AC7577C5C4 CRC64;

## Query Match

Best Local Similarity 47.9%; Score 46; DB 2; Length 284;  
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 HCNFNVDYTRRLRENE 17  
Db 26 HCVPSNITTRRLRELE 40

## RESULT 4

ID 09KER0 PRELIMINARY; PRT; 777 AA.  
AC 09KER0;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DE BH0789 PROTEIN.  
GN BH0789.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=86655;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C-125 / JCM 9153;  
RX MEDLINE=20512582; PubMed=11058132;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331(2000).  
DR EMBL: AP001509; BAB04508.1; -  
KW Complete proteome.  
SQ SEQUENCE 777 AA; 88227 MW; 46BB843DADBEF17B CRC64;

Query Match  
Best Local Similarity 47.9%; Score 46; DB 2; Length 777;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 FEHCNPNVYTRRLRE 15  
Db 136 FKHSNFDYLTQIKE 150

## RESULT 5

ID 093554 PRELIMINARY; PRT; 1307 AA.  
AC 093554; 002539;  
DT 01-FEB-1997 (TREMBLrel. 02, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DE 01-JAN-1999 (TREMBLrel. 09, Last annotation update)  
GN F3A7.5 PROTEIN.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC McMurtry A.;  
RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.  
RN [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
 RA Croxall M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkseen R.,  
 RA Smailson N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL: 281067; CAB02976.1; -  
 SQ SEQUENCE 1307 AA; 147397 MW; EB966398D82D9B9C CRC64;

Query Match 46.9%; Score 45; DB 5; Length 1307;  
 Best Local Similarity 69.2%; Pred. No. 52;  
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 4 HCNFNDVTRLREN 16  
 1 11 11 1111  
 Db 654 CKINDTTRLREN 666

RESULT 6  
 O92CG3 PRELIMINARY; PRT; 205 AA.  
 AC O92CG3;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE MADH DEHYDROGENASE 1 CHAIN J (NUOJ).  
 GN RP790.  
 OS Rickettsia prowazekii.  
 OC Bacteria; Proteobacteriales; alpha subdivision; Rickettsiales;  
 OC Rickettsiaceae; Rickettsiae; Rickettsia.  
 OX NCBI\_TaxID=782;  
 OX 11  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=MADRID E;  
 RC MEDLINE=99039499; PubMed=9823693;  
 RA Anderson S.G.E., Zomrodipour A., Andersson J.O.,  
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;  
 RA Sichteritz-Ponten T., Alsmark U.C.M., Podowski R.M., Neeslund A.K.,  
 RT "The genome sequence of Rickettsia prowazekii and the origin of  
 RT mitochondria.";  
 RL Nature 396:133-140(1998).  
 DR EMBL: AJ235273; CAA15216.1; -  
 DR InterPro: IPR001457; Oxidored\_q3.  
 DR Pfam: PF00499; oxidored\_q3; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 205 AA; 23142 MW; 035BD21245DB276A CRC64;

Query Match 45.8%; Score 44; DB 2; Length 205;  
 Best Local Similarity 66.7%; Pred. No. 12;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 5 HCNFNDVTRLREN 16  
 111 1111111  
 Db 77 HCNFNDVTRLREN 88

RESULT 7  
 O82618 PRELIMINARY; PRT; 513 AA.  
 AC O82618;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE SURFACE GLYCOPROTEIN.  
 GN SPIKE, S1.  
 OS avian infectious bronchitis virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronavirus.  
 OX NCBI\_TaxID=11120;  
 OX 11  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=CONNECTICUT;  
 RC MEDLINE=95159673; PubMed=7656318;  
 RA Wang L., Junker D., Hock L., Edlary E., Collisson E.W.;  
 RT "Evolutionary implications of genetic variations in the S1 gene of  
 RT infectious bronchitis virus.";  
 RL Virus Res. 34:327-338(1994).  
 DR EMBL: L18990; AAA74378.1; -  
 DR InterPro: IPR002551; Corona\_S1.  
 DR Pfam: PF01600; Corona\_S1; 1.  
 SQ SEQUENCE 513 AA; 56541 MW; 2D4C1P986F6B045E CRC64;

Query Match 45.8%; Score 44; DB 12; Length 513;  
 Best Local Similarity 75.0%; Pred. No. 30;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 HCNFNDVT 10  
 11111111  
 Db 97 HCNFSDIT 104

RESULT 8  
 O55343 PRELIMINARY; PRT; 530 AA.  
 AC O55343;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE SPIKE GLYCOPROTEIN S1 SUBUNIT (FRAGMENT).  
 OS avian infectious bronchitis virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronavirus.  
 OX NCBI\_TaxID=11120;  
 OX 11  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=FLORIDA 18288;  
 RC Moore K.M., Jackwood M.W., Bennett J.D., Seal B.S.;  
 RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF027512; AAB87907.1; -  
 DR InterPro: IPR002551; Corona\_S1.  
 DR Pfam: PF01600; Corona\_S1; 1.  
 FT NON\_TER 530 530  
 SQ SEQUENCE 530 AA; 58315 MW; 74CE496CE25688F5 CRC64;

Query Match 45.8%; Score 44; DB 12; Length 530;  
 Best Local Similarity 75.0%; Pred. No. 31;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 HCNFNDVT 10  
 11111111  
 Db 96 HCNFSDIT 103

RESULT 9  
 O27791 PRELIMINARY; PRT; 807 AA.  
 AC O27791;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE COLLAGENASE.  
 GN MTH1763.  
 OS Methanobacterium thermoautotrophicum.

OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
 OC Methanothermobacter.  
 OX NCBI\_TaxID=145262;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DELTA.H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,  
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,  
 RA Spadator R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Wani N., Caruso A., Bush D., Satter H., Patwell D., Prabhakar S.,  
 RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.,  
 RA "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RL J. Bacteriol. 179:7135-7155(1997).  
 DR EMBL: AEO00931; AAB86229.1;  
 DR InterPro: IPR001539; Peptidase\_U32.  
 DR Pfam: PF01136; Peptidase\_U32.  
 DR ProDom: PD004398; Peptidase\_U32; 1.  
 DR PROSITE: PS01276; PEPTIDASE\_U32; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 807 AA; 91152 MW; 466E6FC0F7833D7 CRC64;

Query Match  
 Best Local Similarity 45.8%; Score 44; DB 1; Length 807;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 EHCNDVYTRLRE 15  
 Db 562 FRECDMNDVRLRE 576

RESULT 10  
 Q9ND03 PRELIMINARY; PRT; 1101 AA.  
 AC 09ND03:  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE ATP CITRATE-LYASE.  
 GN CITA-1.  
 OS Clona intestinalis.  
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;  
 OC Clonidae; Clona.  
 OX NCBI\_TaxID=7719;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Hotta K., Takahashi H., Asakura T., Saitoh B., Takatori N., Satou Y.,  
 RA Sato N.;  
 RT "Characterization of Brachyury downstream notochord genes in the Clona  
 RL intestinalis embryo."  
 CC Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO OTHER SUBUNITS (ALPHA) OF SUCCINYL-COA SYNTHETASE.  
 CC -1- SIMILARITY: TO OTHER SUBUNITS (BETA) OF SUCCINYL-COA SYNTHETASE,  
 CC OF MALATE-COA LIGASE AND TO ATP CITRATE-LYASE.  
 DR EMBL: AB036847; BAB00624.1;  
 DR InterPro: IPR000303; COA\_Ligase.  
 DR InterPro: IPR000901; CPase.  
 DR Pfam: PF00549; Ligase\_COA; 1.  
 DR PROSITE: PS00867; CPASE\_2; UNKNOWN\_1.  
 DR PROSITE: PS01216; SUCCINYL-COA\_LIG\_1; 1.  
 DR PROSITE: PS01217; SUCCINYL-COA\_LIG\_2; UNKNOWN\_1.  
 DR Ligase; Lyase; Phosphorylation.  
 SQ SEQUENCE 1101 AA; 120728 MW; 63CFE1B8D285DF49 CRC64;

Query Match  
 Best Local Similarity 45.8%; Score 44; DB 5; Length 1101;  
 Matches 50.0%; Pred. No. 64;

Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 CNDVYTRLRE 15  
 Db 787 CFNELNFKIRE 798

RESULT 11  
 Q82667 PRELIMINARY; PRT; 544 AA.  
 AC 082667:  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE SPIKE PROTEIN SUBUNIT 1 PRECURSOR (FRAGMENT).  
 GN S.  
 OS avian infectious bronchitis virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronavirus.  
 OX NCBI\_TaxID=11120;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-UK/918/67;  
 RX MEDLINE=93298060; PubMed=8390829;  
 RA Cavanagh D., Davis P.J.;  
 RT "Sequence analysis of strains of avian infectious bronchitis  
 RL Arch. Virol. 130:471-476(1992).  
 DR EMBL: X64737; CAA46003.1;  
 DR InterPro: IPR002551; Corona\_S1.  
 DR Pfam: PF01600; Corona\_S1; 1.  
 KW Signal; Membrane.  
 FT CHAIN 1  
 FT NON-TER 19  
 FT SEQUENCE 544 AA; 59856 MW; 0A0A2F26344CE0B CRC64;  
 SQ SEQUENCE 544 AA; 59856 MW; 0A0A2F26344CE0B CRC64;

Query Match  
 Best Local Similarity 44.8%; Score 43; DB 12; Length 544;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 HCNDVDT 10  
 Db 101 HCNDVDT 108

RESULT 12  
 Q82619 PRELIMINARY; PRT; 545 AA.  
 AC 082619:  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE SUPRACE GLYCOPROTEIN (FRAGMENT).  
 GN SPIKE\_S1.  
 OS avian infectious bronchitis virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronavirus.  
 OX NCBI\_TaxID=11120;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HOLTE;  
 RA Wang L., Collisson E.W.;  
 RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: L18988; AAA74379.1;  
 DR InterPro: IPR002551; Corona\_S1.  
 DR Pfam: PF01600; Corona\_S1; 1.  
 FT NON-TER 545  
 FT SEQUENCE 545 AA; 60125 MW; C40BFB969ABA9647 CRC64;

Query Match  
 44.8%; Score 43; DB 12; Length 545;

Best Local Similarity 75.0%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 3 HCNFNDVT 10  
Db 101 HCNFDDIT 108

## RESULT 13

QY9VE6 PRELIMINARY; PRT: 739 AA.  
ID QY9VE6; 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)  
DE 01-JUN-2001 (TREMBlrel. 17, last annotation update)  
CG3136 PROTEIN.  
GN CG3136.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Gelinkler S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Dey A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Egharshian C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Pui Y., Reese M.G.,  
RA Reiter K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Iamos I., Simpson M., Skupski M.P., Smith T.,  
RA Splier E., Spredling A.C., Stapleton M., Strong R., Sun E.,  
RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003787; AAF57345.1; -  
DR FlyBase: FBgn0033010; CG3136.  
DR InterPro: IPR001871; bZIP.  
DR Pfam: PF00170; bZIP.1.  
DR SMART: SM00338; BRLZ.1.  
SO SEQUENCE 739 AA; 84551 MW; 63067E2E08F84F7 CRC64;

Query Match

44.8%; Score 43; DB 5; Length 739;

Best Local Similarity 40.0%; Pred. No. 63;  
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;  
QY 2 EHCNFDVTRLRE 16  
Db 374 KHCNNTTATIKN 388

## RESULT 14

QY9VE6 PRELIMINARY; PRT: 96 AA.  
ID QY9VE6; 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)  
DE 01-JUN-2001 (TREMBlrel. 17, last annotation update)  
ENV. ENVELOPE GLYCOPROTEIN VIA2 REGION (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OC NCBI\_Taxid=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=065;  
RA Cornelissen M., Goudsmit J.;  
RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC MCEVILLY M.M.;  
RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U13544; AAA73733.1; -  
DR InterPro: IPR000777; GP120.  
DR Pfam: PF00516; GP120; 1.  
KW Envelope protein.  
FT NON\_TER 1 1  
FT NON\_TER 96 96  
SQ SEQUENCE 96 AA; 10776 MW; 9B9443D3CD7F43A8 CRC64;

Query Match 44.3%; Score 42.5; DB 12; Length 96;  
Best Local Similarity 43.8%; Pred. No. 10;  
Matches 7; Conservative 6; Mismatches 2; Indels 1; Gaps 1;

QY 2 EHCNFDVTRLRE 17  
Db 43 EYCSFN-ITTEIRDQ 57

## RESULT 15

ID P90081 PRELIMINARY; PRT: 240 AA.  
AC P90081; 01-MAY-1997 (TREMBlrel. 03, Created)  
DT 01-MAY-1997 (TREMBlrel. 03, last sequence update)  
DE 01-JUN-2001 (TREMBlrel. 17, last annotation update)  
ENV. ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OC NCBI\_Taxid=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PATIENT A;  
RX MEDLINE=97184515; PubMed=9032317;  
RA McDonald R.A., Mayers D.L., Chung R.C.Y., Wagner K.F., Kim S.,  
RA Birx D.L., Michael N.L.;  
RT "Evolution of human immunodeficiency virus type 1 env sequence  
cell function.";  
RL J. Virol. 71:1871-1879(1996).  
DR EMBL: U69285; AAC56604.1; -  
DR InterPro: IPR000777; GP120.  
DR Pfam: PF00516; GP120; 1.





NCBI\_TaxID=11676;  
[1]  
SEQUENCE FROM N.A.  
STRAIN-PATIENT A;  
MEDLINE=97184515; PubMed=9032317;  
McDonald R.A., Mayers D.L., Chung R.C.Y., Wagner K.F., Kim S.,  
Blax D.L., Michael N.L.;  
"Evolution of human immunodeficiency virus type 1 env sequence  
variation in patients with diverse rates of disease progression and T-  
cell function.";  
J. Virol. 71:1871-1879(1996).  
EMBL: U69289; AAC56808.1; -;  
DR InterPro: IPR000777; GP120.  
DR Pfam: PF00516; GP120; 1.  
KW Envelope protein.  
FT NON\_TER 1 1  
DR 240 240  
SQ SEQUENCE 240 AA; 26747 MW; B85EA6F749552442 CRC64;

Query Match 44.3%; Score 42.5; DB 12; Length 240;  
Best Local Similarity 57.1%; Pred. No. 25;  
Matches 8; Conservative 5; Mismatches 0; Indels 1; Gaps 1;

QY 2 EHCNFDVTRLRE 15  
Db 32 KNCSPN-ITRLRD 44

RESULT 20  
ID Q9S091 PRELIMINARY; PRT; 313 AA.  
AC Q9S091;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE HYPOTHETICAL 36.0 KDA PROTEIN.  
GN BBO17.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Plasmid cp32-7.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Lathigra R.,  
White O., Dodson R., Hickey E.K., Gwin M., Peterson J., van-Vugt R.,  
Palmer N., Haft D., Rosa P., Stevenson B.;  
"A bacterial genome in flux: The twelve linear and nine circular  
RT extrachromosomal DNAs in an infectious isolate of the Lyme disease  
RT spirochete Borrelia burgdorferi.";  
RL Mol. Microbiol. 0:0-0(1999).  
DR EMBL: AE001579; AAF07613.1; -;  
KW Hypothetical protein; Plasmid.  
SQ SEQUENCE 313 AA; 35968 MW; E81A8779733736C5 CRC64;

Query Match 44.3%; Score 42.5; DB 2; Length 313;  
Best Local Similarity 56.2%; Pred. No. 32;  
Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 4 CNF---NDVTRLREN 16  
Db 211 CNFETNDITTEADEN 226

RESULT 21

ID Q9IK02 PRELIMINARY; PRT; 679 AA.  
AC Q9IK02;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE TRUNCATED ENVELOPE GLYCOPROTEIN.

ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=546HC-F7;  
RC Gartner S., Liu Y., Tang X.P., McArthur J.C., Scott J.;  
RT "Analysis of human immunodeficiency virus type 1 gp160 sequences from  
a patient with HIV dementia: evidence for monocyte trafficking into  
RT brain.";  
J. Neurovirol. 0:0-0(2000).  
EMBL: AF217165; AAF7507.1; -;  
DR InterPro: IPR000328; Env\_GP41.  
DR InterPro: IPR000777; GP120.  
DR Pfam: PF00516; GP120; 1.  
DR Pfam: PF00517; GP41; 1.  
KW Envelope protein.  
SQ SEQUENCE 679 AA; 76494 MW; 69C6485D07902578 CRC64;

Query Match 44.3%; Score 42.5; DB 12; Length 679;  
Best Local Similarity 57.1%; Pred. No. 70;  
Matches 8; Conservative 5; Mismatches 0; Indels 1; Gaps 1;

QY 2 EHCNFDVTRLRE 15  
Db 153 KNCSPN-ITRLRD 165

RESULT 22  
ID Q9IK06 PRELIMINARY; PRT; 684 AA.  
AC Q9IK06;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN.  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=546HC-B6;  
RA Gartner S., Liu Y., Tang X.P., McArthur J.C., Scott J.;  
RT "Analysis of human immunodeficiency virus type 1 gp160 sequences from  
a patient with HIV dementia: evidence for monocyte trafficking into  
RT brain.";  
J. Neurovirol. 0:0-0(2000).  
EMBL: AF217161; AAF75503.1; -;  
DR InterPro: IPR000328; Env\_GP41.  
DR InterPro: IPR000777; GP120.  
DR Pfam: PF00516; GP120; 1.  
DR Pfam: PF00517; GP41; 1.  
SQ SEQUENCE 684 AA; 77266 MW; F9E6F3CCE4D32E10 CRC64;

Query Match 44.3%; Score 42.5; DB 12; Length 684;  
Best Local Similarity 57.1%; Pred. No. 70;  
Matches 8; Conservative 5; Mismatches 0; Indels 1; Gaps 1;

QY 2 EHCNFDVTRLRE 15  
Db 153 KNCSPN-ITRLRD 165

RESULT 23  
ID Q73350 PRELIMINARY; PRT; 852 AA.  
AC Q73350;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-JUN-2001 (TREMblrel. 17, last annotation update)  
 DE ENVELOPE GLYCOPROTEIN.  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BU/91/07;  
 RX MEDLINE=96303593; PubMed=8744585;  
 RA Penny M.A., Thomas S.J., Douglas N.W., Ranjbar S., Holmes H.,  
 RA Daniels R.S.;  
 RT "env gene sequences of primary HIV type 1 isolates of subtypes B, C,  
 RT D, E, and F obtained from the World Health Organization Network for  
 RT HIV Isolation and Characterization."  
 RL AIDS Res. Hum. Retroviruses 12:741-747(1996).  
 DR EMBL; U39245; AAB37182.1; -  
 DR InterPro; IPR000328; Env.GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 RW Envelope protein.  
 SQ SEQUENCE 852 AA; 96383 MW; 8F2AC94B5FE7492B CRC64;

Query Match 44.3%; Score 42.5; DB 12; Length 852;  
 Best Local Similarity 50.0%; Pred. No. 87;  
 Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

OY 2 EHCNENDVTRLRE 17  
 Db 155 KNCSEF-ITTRLRD 165

RESULT 24  
 OYK05 PRELIMINARY; PRT; 861 AA.  
 AC OYK05;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE ENVELOPE GLYCOPROTEIN.  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=546HC-C5;  
 RA Gartner S., Liu Y., Tang X.P., McArthur J.C., Scott J.;  
 RT "Analysis of human immunodeficiency virus type 1 gp160 sequences from  
 RT brain."  
 RT J. Neurovirol. 0:0-0(2000).  
 DR EMBL; AF217163; AAF75504.1; -  
 DR InterPro; IPR000328; Env.GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 SQ SEQUENCE 861 AA; 97808 MW; 02972C8F2C9464D3 CRC64;

Query Match 44.3%; Score 42.5; DB 12; Length 861;  
 Best Local Similarity 57.1%; Pred. No. 88;  
 Matches 8; Conservative 5; Mismatches 0; Indels 1; Gaps 1;

OY 2 EHCNENDVTRLRE 15  
 Db 153 KNCSEF-ITTRLRD 165

RESULT 25  
 OYK04

ID OYK04 PRELIMINARY; PRT; 861 AA.  
 AC OYK04;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE ENVELOPE GLYCOPROTEIN.  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=546HC-C8;  
 RA Gartner S., Liu Y., Tang X.P., McArthur J.C., Scott J.;  
 RT "Analysis of human immunodeficiency virus type 1 gp160 sequences from  
 RT brain."  
 RT J. Neurovirol. 0:0-0(2000).  
 DR EMBL; AF217163; AAF75505.1; -  
 DR InterPro; IPR000328; Env.GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 SQ SEQUENCE 861 AA; 97763 MW; 0CE8C39E698D375 CRC64;

Query Match 44.3%; Score 42.5; DB 12; Length 861;  
 Best Local Similarity 57.1%; Pred. No. 88;  
 Matches 8; Conservative 5; Mismatches 0; Indels 1; Gaps 1;

OY 2 EHCNENDVTRLRE 15  
 Db 153 KNCSEF-ITTRLRD 165

Search completed: April 2, 2002, 09:23:52  
 Job time: 196 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 10:43:55 ; Search time 23.72 Seconds

(Without alignments)  
53.088 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58

Perfect score: 96

Sequence: 1 FEHCNFNDVTTRLRENE 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database :  
1: A\_Geneseq\_1101.\*  
2: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
3: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
4: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
5: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
6: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
7: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
8: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
9: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
10: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
11: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
12: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
13: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
14: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
15: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
16: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
17: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
18: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
19: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
20: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
21: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
22: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
23: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	77	20	AAV27313
2	96	100.0	102	18	AAW09041
3	96	100.0	103	12	AA11426
4	96	100.0	103	14	AA132291
5	96	100.0	107	16	AA180239
6	96	100.0	115	21	AA180364
7	96	100.0	115	21	AA180376
8	96	100.0	115	21	AA180376
9	96	100.0	127	20	AAV27311
10	96	100.0	128	11	AA1804704
11	96	100.0	128	11	AA1804744

12	96	100.0	128	12	AA11876	Human lymphocyte s
13	96	100.0	128	16	AA180240	Human membrane att
14	96	100.0	128	16	AA180315	Human CD59. Homo
15	96	100.0	128	18	AA126318	Human CD59. Homo
16	96	100.0	128	22	AA100688	Human CD59 protein
17	96	100.0	260	22	AA100685	Human CTLA4-human
18	96	100.0	261	22	AA100684	Peptide #4266 enco
19	68	70.8	58	22	AA117832	Peptide #4376 enco
20	68	70.8	58	22	AA103039	Peptide #4161 enco
21	68	70.8	58	22	AA105479	Baboon CD59 protel
22	68	70.8	75	20	AAV27314	African green monk
23	63	65.6	77	20	AA127315	African green monk
24	63	65.6	126	16	AA186964	Ly-6 terminal CIP
25	63	65.6	126	16	AA186310	Baboon complement
26	63	65.6	128	16	AA186983	Ly-6 terminal CIP
27	63	65.6	128	16	AA186309	Rabbit CD59 protel
28	49	51.0	77	20	AAV27321	Owl monkey CD59 pr
29	49	51.0	77	20	AAV27316	Rabbit CD59 protel
30	49	51.0	124	20	AAV27312	Owl monkey complem
31	49	51.0	128	16	AA186986	Ly-6 terminal CIP
32	49	51.0	128	16	AA186312	Marmoset CD59 prot
33	47	49.0	77	20	AAV27317	Marmoset complement
34	47	49.0	128	16	AA186987	Ly-6 terminal CIP
35	47	49.0	128	16	AA186313	Arabidopsis thalia
36	41	42.7	144	21	AA186378	Arabidopsis thalia
37	41	42.7	177	21	AA186377	Arabidopsis thalia
38	41	42.7	501	22	AA181996	Amino acid sequenc
39	41	42.7	501	22	AA181999	Amino acid sequenc
40	41	42.7	536	8	AA170142	Sequence of a regl
41	41	42.7	537	19	AA14943	Avian infectious b
42	41	42.7	1093	14	AA180815	Truncated IBV splk
43	41	42.7	1162	7	AA160720	Sequence of the S1
44	41	42.7	1162	8	AA170137	Sequence of a regl
45	41	42.7	1162	17	AA106829	Major neutralising
46	41	42.7	1644	22	AA182000	Amino acid sequenc
47	41	42.7	1645	22	AA181997	Amino acid sequenc
48	41	42.7	3433	18	AA122017	Utrrophin. Homo sa
49	40	41.7	217	17	AA189325	Pilin protein. Ha
50	40	41.7	403	21	AA130662	Arabidopsis thalia

#### ALIGNMENTS

RESULT 1	AAV27313	standard; peptide: 77 AA.
ID	AAV27313	
XX	AAV27313:	
AC	AAV27313:	
XX	05-NOV-1999 (first entry)	
DT	XX	
XX	Human CD59 protein fragment.	
XX	CD59 mediated complement; human; Cd59 protein; mimetic;	
KW	tumour therapy; complement-mediated inflammation; immune disorder;	
KW	immunovascularitis; rheumatoid arthritis; scleroderma; C5b-9 complex;	
KW	plasma membrane antigen.	
XX		
OS	Homo sapiens.	
XX		
PN	MO9940115-A2.	
XX		
PD	12-AUG-1999.	
XX		
PF	09-FEB-1999; 99WO-US02802.	
XX		
PR	09-FEB-1998; 98US-0020393.	
XX		
PA	(BLOO-) BLOOD CENT RES FOUND INC.	
PA	(OKLA-) OKLAHOMA MEDICAL RES FOUND.	
XX		
PI	Sims PJ;	

XX  
DR WPI: 1999-527301/44.  
XX  
PI Compounds modulating CD59 mediated complement activity, for  
PI treatment of, e.g. immunovascularitis  
XX  
PS Disclosure: Fig 8A-B: 75pp; English.  
XX  
CC The invention relates to compounds modulating CD59 mediated complement  
CC activity. It provides (1) molecules structurally mimicking human CD59  
CC amino acid residues 42-58 (region which serves as binding site for CD59  
CC -C9 interactions) when they are in a spatial orientation which can  
CC inhibit the formation of the human C5b-9 complex. These mimetics  
CC specifically bind to amino acid residues 359-384 of human C9. (11)  
CC in a spatial orientation mimicking C9 amino acids 359-384 when they are  
CC complex. Compounds that mimic CD59 can be used to increase CD59  
CC inhibition of C5b-9 complex assembly. This is especially useful in  
CC patients in need of suppression of complement-mediated inflammation, e.g.  
CC arthritis, scleroderma, and diseases such as immunovascularitis, rheumatoid  
CC C5b-9 complex assembly. This is useful in patients in need of promote  
CC activation. The present sequence represents a human CD59 (a plasma membrane  
CC antigen) protein fragment.  
CC  
XX  
SO Sequence 77 AA;

Query Match  
Best Local Similarity 100.0%; Score 96; DB 20; Length 77;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHCNENDVYTRIRENE 17  
DB 42 fchcnndvtrlirene 58  
|||||

RESULT 2  
ID AAM09041  
AAW09041 standard; Protein; 102 AA.  
AC AAM09041;  
XX  
DT 23-SEP-1997 (first entry)  
XX  
DE Human membrane attack complex inhibitory factor mutant Asn18Gln.  
XX  
KW Human; membrane attack complex; inhibitory factor; MACIF; mutant;  
KW fibroblast growth factor; mediation; cytolysis; PG12; bFGF; basic;  
XX transplant rejection.  
OS Homo sapiens.  
XX  
FH Key  
FT Peptide 1..25 Location/Qualifiers  
FT Peptide 26..102 /label= sig\_peptide  
FT /label= mat\_peptide  
FT Misc-difference 43 /note= "wild type Asn replaced by Gln"  
XX  
PN W09700320-A1.  
XX  
PD 03-JAN-1997.  
XX  
PF 13-JUN-1996; 96MO-JP01609.  
XX  
PR 16-JUN-1995; 95JP-0174282.  
XX  
PA (YAMA ) YAMANOCHI PHARM CO LTD.  
XX

PI Egashira A, Itou K, Masuho Y, Sugita Y, Suzuki H;  
PI Yamaji N, Yasunaga K;  
XX  
DR WPI: 1997-077527/07.  
DR N-PSDB; AAT49584.  
XX  
PT Modified human membrane attack complex inhibitory factor - has  
PT glutamine at position 18 and is an antiinflammatory and complement  
PT blocker, useful for inhibition of transplant hyper-acute rejection  
XX  
PS Claim 1; Page 35; 59pp; Japanese.  
XX  
CC The present sequence is the human membrane attack complex  
CC inhibitory factor (MACIF) mutant Asn18Gln (residues 1-77), which  
CC basic fibroblast growth factor (bFGF) by complement. It can be used  
CC to suppress transplanted organ rejection, especially hyperacute  
CC transplant rejection, and the non-lethal effects of complement,  
CC e.g. inflammation.  
XX  
SO Sequence 102 AA;

Query Match  
Best Local Similarity 100.0%; Score 96; DB 18; Length 102;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHCNENDVYTRIRENE 17  
DB 67 fchcnndvtrlirene 83  
|||||

RESULT 3  
ID AAR11426  
AAR11426 standard; Protein; 103 AA.  
XX  
AC AAR11426;  
XX  
DT 11-JUN-1991 (first entry)  
XX  
DE Human lymphocyte surface antigen.  
XX  
KW Lymphocyte; surface antigen; autoimmune disease; cancer.  
XX  
OS Homo sapiens.  
XX  
PN JP03048696-A.  
XX  
PD 01-MAR-1991.  
XX  
PF 14-JUL-1989; 89JP-0183264.  
XX  
PR 14-JUL-1989; 89JP-0183264.  
XX  
PA (TORA ) TORAY IND INC.  
XX  
DR WPI: 1991-106290/15.  
DR N-PSDB; AAQ11251.  
XX  
PT Surface antigen of human lymphocyte - for use in study of human  
PT immune system of 101 residues  
XX  
PS Claim 1; Page 1; 12pp; Japanese.  
XX  
CC This protein is similar to the murine lymphocyte antigen Ly 6.  
CC Antibodies raised against the antigen can be used in formulations  
CC for the treatment of autoimmune diseases, cancer and infections.  
CC The corresponding nucleotide sequence does not include an initial  
CC leu codon.  
XX  
SO See also AAQ11252-4.  
XX  
SO Sequence 103 AA;

```

Query Match      100.0%; Score 96; DB 12; Length 103;
Best Local Similarity 100.0%; Pred. No. 7.2e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNENDVTRLRRENE 17
   |||||
Db 42 fenhcnfndvtrlrrene 58

RESULT 4
AAR32291
ID AAR32291 standard; Protein; 103 AA.
AC AAR32291;
XX
XX 08-JUN-1993 (first entry)
DE Sequence of CD59.
XX
XX Complement mediated attack inhibitor; CD59; transplant rejection; ss.
XX
XX Homo sapiens.
XX
XX WO9302188-A.
XX
XX 04-FEB-1993.
XX
XX 14-JUL-1992; 92MO-US05920.
XX
XX 15-JUL-1991; 91US-0729926.
XX
XX 29-JUN-1992; 92US-0906394.
XX
XX (OKLA-) OKLAHOMA MED RES FOUND.
XX
XX (UYVA ) UNIV YALE.
XX
XX Bell L, Bothwell ALM, Elliot EA, Flavell RA, Madri J;
XX
XX Rollins S, Sims PJ, Squinto S;
XX
XX WPI: 1993-058786/07.
XX
XX N-PSDB; AAQ36708.
XX
XX Genetically engineered mammalian cell for treatment of coronary
XX
XX artery disease - inhibits complement-mediated attack and does not
XX
XX express surface proteins encoded by class I or II major
XX
XX histocompatibility complex genes
XX
XX Claim 31: Page 75; 89pp; English.
XX
XX The inventors claim a cell which contains a gene sequence which
XX
XX encodes protein CD59. CD59 is expressed by the cell and CD59 inhibits
XX
XX complement mediated attack of the cell. The cells fail to elicit T
XX
XX lymphocyte mediated attack or are resistant to complement mediated
XX
XX attack. They can be used to treat patients with immune disorders.
XX
XX Sequence 103 AA:

Query Match      100.0%; Score 96; DB 14; Length 103;
Best Local Similarity 100.0%; Pred. No. 7.2e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNENDVTRLRRENE 17
   |||||
Db 42 fenhcnfndvtrlrrene 58

RESULT 5
AAR80239
ID AAR80239 standard; Protein; 107 AA.
AC AAR80239;
XX
XX

```

```

DT 17-APR-1996 (first entry)
XX
XX Human membrane attack complex inhibition factor.
DE
XX
XX MACIF; membrane attack complex inhibition factor; complement system;
XX
XX regulation; activation; final stage; inhibit damage; disease therapy;
XX
XX type II allergy; type III allergy; inflammatory disease treatment;
XX
XX phosphatidylinositol anchor; glycoprotein; diagnosis.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX
XX Peptide 1..25
XX
XX Peptide /label= secretory_signal_sequence
XX
XX /note= "optionally not present, replaced with
XX
XX H or Met"
XX
XX Peptide 101..107
XX
XX /note= "optionally not present"
XX
XX Peptide 102..107
XX
XX /note= "optionally not present"
XX
XX Protein 26..95
XX
XX /note= "core human MACIF protein"
XX
XX Peptide 96..101
XX
XX /label= PI_attachment_signal_sequence
XX
XX Modified-site 101
XX
XX /label= PI_anchor
XX
XX /note= "modified by PI - skeletal structure composed
XX
XX of phospho-ethanolamine, glycan and
XX
XX phosphatidylinositol"
XX
XX
XX EP672683-A1.
XX
XX 20-SEP-1995.
XX
XX 19-APR-1990; 90EP-0200379.
XX
XX
XX 27-OCT-1989; 89JP-0281197.
XX
XX 21-APR-1989; 89JP-0103088.
XX
XX 12-JUL-1989; 89JP-0179933.
XX
XX 06-SEP-1989; 89JP-0230993.
XX
XX 13-SEP-1989; 89JP-0238246.
XX
XX 21-SEP-1989; 89JP-0247818.
XX
XX
XX (YAMA ) YAMANOUCHI PHARM CO LTD.
XX
XX Furuchi K, Ito K, Sugita Y, Takayama M, Tomita M;
XX
XX Yamaji N, Yano S, Yusakawa K, Kasukawa K, Takemoto T;
XX
XX WPI: 1995-321975/42.
XX
XX N-PSDB; AAQ98532.
XX
XX Peptide with human membrane attack complex inhibition factor
XX
XX activity - also DNA and expression vectors used to regulate the
XX
XX complement system in the final stage of complement activation
XX
XX Claim 2; Page 28; 49pp; English.
XX
XX Human membrane attack complex inhibition factor (MACIF) regulates the
XX
XX complement system in the final stage of complement activation, and
XX
XX inhibits damage of human cells and tissues as a result of MAC
XX
XX formation. Naturally occurring human MACIF is a glycoprotein of mol. wt.
XX
XX 18 +/- 1 kDa (by SDS-PAGE) with a phosphatidylinositol (PI) anchor at
XX
XX position 76 (Glu) at the C-terminus. When the gene encoding MACIF is
XX
XX expressed in bacteria, the gene (see AAQ98532) gives a modified human
XX
XX MACIF protein comprising 128 amino acid residues (AAR80240); the PI
XX
XX anchor attachment does not occur in that case.
XX
XX Sequence 107 AA:

Query Match      100.0%; Score 96; DB 16; Length 107;
Best Local Similarity 100.0%; Pred. No. 7.5e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 FEHCNENDVTTLRLENE 17  
 DB 67 fehcnfndvttlrleene 83

## RESULT 6

AAAG03764  
 ID AAG03764 standard; Protein; 115 AA.

AC AAG03764;

DT 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 7845.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 gene therapy; chromosome mapping.

OS Homo sapiens.

PN EP1033401-A2.

PD 06-SEP-2000.

PE 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

PA (GEST ) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI: 2000-500381/45.

DR N-PSDB; AAC03770.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 PS Claim 13; SEQ ID 7845; 71bp + CD-ROM; English.

CC The present sequence is a polypeptide encoded by one of a large number  
 of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 were prepared from total human RNAs or polyA+ RNAs derived from 30  
 different tissues. EST sequences usually correspond mainly to the 3'  
 untranslated region (UTR) of the mRNA because they are often obtained  
 from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 those cases where longer cDNA sequences have been obtained, the full 5'  
 UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 ends and can therefore be used to obtain full length cDNAs and genomic  
 DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 chromosome mapping procedures. They are used to obtain upstream  
 regulatory sequences and to design expression and secretion vectors.

SO Sequence 115 AA;

Query Match 100.0%; Score 96; DB 21; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-08;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FEHCNENDVTTLRLENE 17  
 DB 67 fehcnfndvttlrleene 83

## RESULT 7

AAAG3765

ID AAG03765 standard; Protein; 115 AA.

AC AAG03765;

XX 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 7846.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 gene therapy; chromosome mapping.

OS Homo sapiens.

PN EP1033401-A2.

PD 06-SEP-2000.

PE 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

PA (GEST ) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI: 2000-500381/45.

DR N-PSDB; AAC03771.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 PS Claim 13; SEQ ID 7846; 71bp + CD-ROM; English.

CC The present sequence is a polypeptide encoded by one of a large number  
 of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 were prepared from total human RNAs or polyA+ RNAs derived from 30  
 different tissues. EST sequences usually correspond mainly to the 3'  
 untranslated region (UTR) of the mRNA because they are often obtained  
 from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 those cases where longer cDNA sequences have been obtained, the full 5'  
 UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 ends and can therefore be used to obtain full length cDNAs and genomic  
 DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 chromosome mapping procedures. They are used to obtain upstream  
 regulatory sequences and to design expression and secretion vectors.

SO Sequence 115 AA;

Query Match 100.0%; Score 96; DB 21; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-08;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FEHCNENDVTTLRLENE 17  
 DB 67 fehcnfndvttlrleene 83

## RESULT 8

AAAG3766

ID AAG03766 standard; Protein; 115 AA.

AC AAG03766;

DT 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 7847.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 gene therapy; chromosome mapping.

OS Homo sapiens.

PN EP1033401-A2.



```

XX 06-SEP-2000.
PD
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX
XX (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
DR WPI; 2000-500381/45.
XX
DR N-PSDB; AAC03772.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 13; SEQ ID 7847; 71pp + CD-ROM; English.
XX
XX The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNAs or polyA+ RNAs derived from 30'
CC different tissues. EST sequences usually correspond mainly to the 3'
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.
XX
XX Sequence 115 AA;
SO

Query Match 100.0%; Score 96; DB 21; Length 115;
Best Local Similarity 100.0%; Pred. No. 8.1e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FEHCNFNDVTRRLRENE 17
| | | | | | | | | | | | | | |
Db 67 fchcnfndvtrrlrene 83

RESULT 9
AY27311
ID AY27311 standard; peptide: 127 AA.
AC AY27311;
XX
XX 05-NOV-1999 (first entry)
DT
XX
XX Human CD59 protein sequence.
DE
XX
XX CD59 mediated complement; human; Cd59 protein; C9 protein; mimetic;
KW tumour therapy; complement-mediated inflammation; immune disorder;
KW immunovascularitis; rheumatoid arthritis; scleroderma; C5b-9 complex;
KW plasma membrane antigen.
XX
XX Homo sapiens.
OS
XX
XX MO9940115-A2.
PN
XX 12-AUG-1999.
PD
XX
XX 09-FEB-1999; 99WO-US02802.
PF
XX
XX 09-FEB-1998; 98US-0020393.
PR
XX
XX (BLOO-) BLOOD CENT RES FOUND INC.
XX
XX (OKLA-) OKLAHOMA MEDICAL RES FOUND
PA

```

XX	Sims PJ;
PI	
XX	
DR	WPI; 1999-527301/44.
XX	
PT	Compounds modulating CD59 mediated complement activity, for
PT	treatment of, e.g. Immunovasculitis
XX	
PS	Disclosure; Fig 1A; 75pp; English.
XX	
CC	The invention relates to compounds modulating CD59 mediated complement
CC	activity. It provides (i) molecules structurally mimicking human CD59
CC	amino acid residues 42-58 (region which serves as binding site for CD59
CC	-C9 interactions) when they are in a spatial orientation which can
CC	inhibit the formation of the human C5b-9 complex. These mimetics
CC	specifically bind to amino acid residues 359-384 of human C9. (ii)
CC	molecules structurally mimicking C9 amino acids 359-384 when they are
CC	in a spatial orientation which can promote the formation of the C5b-9
CC	complex. Compounds that mimic CD59 can be used to increase CD59
CC	inhibition of C5b-9 complex assembly. This is especially useful in
CC	patients in need of suppression of complement-mediated inflammation, e.g.
CC	immune disorders and diseases such as immunovasculitis, rheumatoid
CC	arthritis, scleroderma. Compounds that mimic C9 can be used to promote
CC	C5b-9 complex assembly. This is useful in patients in need of complement
CC	activation. The composition can be administered as an adjunct to tumour
CC	therapy. The present sequence represents a human CD59 (a plasma membrane
CC	antigen) protein sequence.
XX	
SQ	Sequence 127 AA:
OY	1 PEHCNFDVTTFLRENE 17                               66 fchcnfdvttflrene 82
DB	
RESULT 10	
AAR04704	
ID	AAR04704 standard; protein; 128 AA.
XX	
AC	AAR04704;
XX	
DT	30-AUG-1990 (first entry)
XX	
DE	Sequence of the I F5 antigen derived from human cell membrane.
XX	
KW	Human IF5 antigen; Immunodiagnosis; pernicious anaemia; rheumatoid arthritis; systemic lupus erythematosus; glomerular nephritis.
OS	Homo sapiens.
PN	EP351313-A.
PD	17-JAN-1990.
XX	
PF	11-JUL-1989; 89EP-040196.
XX	
PR	23-MAY-1989; 89JP-0172187.
PA	(MITU ) MITSUBISHI KASEI CORP.
XX	
PI	Okada H, Okada N, Nagami Y, Takahashi K, Takizawa H, Kondo J;
XX	
DR	WPI; 1990-016630/03.
DR	N-PSDB; AAO03116, AAN93318.
XX	
PT	New glyco:protein IF5 antigen -
PT	derived from human cell membrane, inhibits complement-mediated
PT	cell membrane damage

XX Claim 3; p. 11; 26pp; English.  
PS  
XX

CC Typically it is prepd. from human erythrocytes which are centrifuged and  
CC the cell membrane fraction suspended overnight in buffer contg. 18  
CC n-octyl-beta-D-glucopyranoside (MOG). After centrifugation, the super-  
CC nant is treated with solid (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> to 60% satn. After centrifugation  
CC the ppte. is dissolved in buffer contg. 0.1%NOG, then dialysed overnight  
CC against the same plus 0.15 M NaCl. 1F5 antigen is a glycoprotein  
CC with a mol. wt. of 20-25 kd. It contains N-glycoside type carbohydrate  
CC chain and phosphatidylinositol. It inhibits complement-mediated cell  
CC membrane damage. It may be used to generate polyclonal or  
CC monoclonal antibodies (abs) which may be used to determine  
CC the presence of 1F5 antigen on the surface of erythrocytes,  
CC lymphocytes or other cells, thus enabling the diagnosis of diseases,  
CC such as pernicious anaemia, rheumatoid arthritis,  
CC systemic lupus erythematosus and glomerular nephritis. It may be used  
CC to treat disorders in which complement activation is involved. Abs to it  
CC may be used in targeting therapy, such as cancer cell lysis, or  
CC elimination of malignant cells. Residues 27-70 are encoded by  
CC a cDNA fragment of the gene which encodes 1F5 antigen prepd. by  
CC the polymerase chain reaction (in AAN93318).  
XX  
SQ Sequence 128 AA;

Query Match 100.0%; Score 96; DB 11; Length 128;  
Best Local Similarity 100.0%; Pred. No. 9e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 FEHCNFDVYTRLRENE 17  
Db 67 fehcnfdvtrlrrene 83

RESULT 11  
AAR07444  
ID AAR07444 standard; protein; 128 AA.  
XX  
AC AAR07444;

XX 28-JAN-1991 (first entry)  
DT XX  
XX  
DE Human membrane attack complex inhibition factor (MACIF) gene  
DE product.  
XX  
XX  
KW Haemolysis; late complement components.  
XX  
OS Homo sapiens.  
OS  
FH Key  
FH Location/Qualifiers  
FT Protein 26..128  
FT /label=Mature MACIF protein.  
FN EP94035-A.  
XX  
XX  
XX 24-OCT-1990.  
PD  
XX  
XX 19-APR-1990; 90EP-0304203.  
PF  
XX

PR 27-OCT-1989; 89JP-0281197.  
PR 21-APR-1989; 89JP-0103088.  
PR 12-JUL-1989; 89JP-0179933.  
PR 06-SEP-1989; 89JP-0230983.  
PR 13-SEP-1989; 89JP-0238246.  
PR 21-SEP-1989; 89JP-0247818.  
XX  
XX

PA (YAMA ) YAMANOUCHI PHARM KK.  
XX

PI Tomita M, Sugita Y, Takemoto T, Furuchi K, Takayama M;  
PI Tsuchikawa K, Yano S, Yamaji N, Ito K;  
XX  
XX WPI; 1990-322496/43.

DR N-PSDB; AA006262.  
XX

PT Genes encoding protein with human MACIF activity - also  
PT expression vectors and proteins produced from expression of the  
PT genes.  
XX  
XX

PS Disclosure; Fig 1; 49pp; English.  
XX

CC Gene product may be expressed in large quantities and pure form  
CC from CHO cells, useful for inhibiting the activity of late  
CC complement components ie. haemolysis resulting from MAC formation.  
XX  
SQ Sequence 128 AA;

Query Match 100.0%; Score 96; DB 11; Length 128;  
Best Local Similarity 100.0%; Pred. No. 9e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 FEHCNFDVYTRLRENE 17  
Db 67 fehcnfdvtrlrrene 83

RESULT 12  
AAR11876  
ID AAR11876 standard; Protein; 128 AA.  
XX  
AC AAR11876;

XX 24-JUL-1991 (first entry)  
DT XX  
XX  
DE Human lymphocyte surface antigen precursor.  
DE  
XX  
KW Ly-6; cancer; autoimmune disease; pre A precursor.  
XX  
OS Homo sapiens.  
OS  
FH Key  
FH Location/Qualifiers  
FT Peptide 1..25  
FT /label= Pre A precursor  
FT Protein 26..128  
FT /label= Mature surface antigen  
PN JP03081297-A.  
XX  
XX  
PD 05-APR-1991.  
XX  
XX 23-AUG-1989; 89JP-0218183.  
PF  
XX  
XX 23-AUG-1989; 89JP-0218183.  
PR  
XX  
XX (TORA ) TORAY IND INC.  
PA  
XX  
XX  
XX WPI; 1991-144848/20.  
DR N-PSDB; AAQ11684-6.  
XX  
XX

PT Surface antigen of human lymphocyte - and gene and antibody  
PT useful for investigation and treatment of autoimmune disease,  
PT cancer, infection, etc.  
XX  
XX

PS Claim 5; Page 1203; 16pp; Japanese.  
XX

CC Clone is derived from the Ly-6 cDNA sequence from a human lymphocyte  
CC cDNA library. The product may be useful in the study and development  
CC of drugs for the treatment of autoimmune diseases and cancer. Probes  
CC may also be developed for the isolation of other gene families.  
XX  
SQ Sequence 128 AA;

Query Match 100.0%; Score 96; DB 12; Length 128;  
Best Local Similarity 100.0%; Pred. No. 9e-08;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 FEHCNFNDVTRIRENE 17  
| | | | | | | | | | | | | | | | | | | | |  
Db 67 fehcnfndvtrirene 83

RESULT 13  
AAR80240

ID AAR80240 standard; Protein: 128 AA.

AC AAR80240;

DT 17-APR-1996 (first entry)

DE Human membrane attack complex inhibition factor.

KW MAC1F, membrane attack complex inhibition factor; complement system;

KM regulation; activation; final stage; inhibit damage; disease therapy;

KW type II allergy; type III allergy; inflammatory disease treatment;

KM phosphatidylinositol anchor; glycoprotein; diagnosis.

OS Homo sapiens.

FH Key

FT Peptide

FT Protein

FT Binding-site

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

Location/Qualifiers  
1..25  
/label= secretory-signal\_sequence  
26..128  
/note="human MAC1F"  
96..128  
/label= PI\_anchor\_attachment\_signal\_sequence

EP672683-A1.

20-SEP-1995.

19-APR-1990;

90EP-0200379.

27-OCT-1989;

89JP-0281197.

21-APR-1989;

89JP-0103088.

12-JUL-1989;

89JP-0179933.

06-SEP-1989;

89JP-0230983.

13-SEP-1989;

89JP-0238246.

21-SEP-1989;

89JP-0247818.

(YAMA ) YAMANOUCHI PHARM CO LTD.

Furuichi K, Ito K, Sugita Y, Takayama M, Tomita M;

Yamaji N, Yano S, Yusakawa K, Kasukawa K, Takemoto T;

WPI: 1995-321975/42.

N-PSDB: AAO98532.

Peptide with human membrane attack complex inhibition factor

activity - also DNA and expression vectors used to regulate the

complement system in the final stage of complement activation

Disclosure; Fig 2; 49pp; English.

Human membrane attack complex inhibition factor (MAC1F) regulates the

complement system in the final stage of complement activation, and

inhibits damage of human cells and tissues as a result of MAC

formation. Naturally occurring human MAC1F is a glycoprotein of mol. wt.

18 +/- 1 kDa (by SDS-PAGE) with a phosphatidylinositol (PI) anchor at

position 76 (Glu) at the C-terminus. When the gene encoding MAC1F is

expressed in bacteria, the gene (see AAO98532) gives a modified human

MAC1F protein comprising 128 amino acid residues (AAR80240); the PI

anchor attachment does not occur in that case.

Sequence 128 AA;

Query Match 100.0%; Score 96; DB 16; Length 128;

Best Local Similarity 100.0%; Pred. No. 9e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 FEHCNFNDVTRIRENE 17  
| | | | | | | | | | | | | | | | | | | | |  
Db 67 fehcnfndvtrirene 83

RESULT 14  
AAR86315

ID AAR86315 standard; Protein: 128 AA.

AC AAR86315;

DT 11-MAR-1996 (first entry)

DE Human CD59.

KW Terminal complement inhibitor protein; terminal CIP; CD59; CD46;

KM transmembrane terminal CIP; TWTCIP; Ly6 antigen; transgenic animal;

KM organ transplantation; glycosyl-phosphatidylinositol; GPI.

OS Homo sapiens.

FH Key

FT Peptide

FT Region

FT Region

FT Region

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

Location/Qualifiers  
1..25  
/label= sig-peptide  
26..96  
/label= Ly6 motif  
103..128  
/label= GPI\_anchor\_region  
/note="GPI anchor region corresponds to  
amino acids 78-103 of the mature  
CD59 protein"

WO9523512-A1.

08-SEP-1995.

01-MAR-1995;

95WO-US02944.

03-MAR-1994;

94US-0205720.

(ALEX-) ALEXION PHARM INC.

Rollins S, Rother RP, Squinto SP;

WPI: 1995-320335/41.

N-PSDB: AAT03338.

Terminal complement inhibitor chimeric protein and nucleic acid -

esp. against human complement, useful for protecting cells from

complement attack e.g. in organ transplantation

Disclosure; Page 63-64; 85pp; English.

Human CD59 (AAR86315) is a terminal complement inhibitor protein (CIP)

that includes a C-terminal region involved in directing attachment of

a glycosyl-phosphatidylinositol (GPI) anchor. A chimeric transmembrane

terminal CIP (TWTCIP) comprising amino acids 1-77 of mature CD59

(i.e. lacking the GPI anchor region) and the transmembrane domain

of amino acids 270-350 of human CD46, a membrane cofactor protein.

The TWTCIP has been expressed on the cell surfaces of the organs of

transgenic animals. Such transgenic organs are protected from human

complement attack upon transplantation.

Sequence 128 AA;

Query Match 100.0%; Score 96; DB 16; Length 128;

Best Local Similarity 100.0%; Pred. No. 9e-08;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHCNFNDVTRRLRENE 17  
 Db 67 fehcnfndvtrrlrene 83

RESULT 15  
 AAM26318  
 ID AAM26318 standard; Protein: 128 AA.  
 AC AAM26318;  
 XX  
 DT 17-NOV-1997 (first entry)  
 DE Human CD59.  
 KW CD59; MAC1F; protectin; p18; human; C5b-9 complement inhibitor;  
 KM retrovirus; vector; gene therapy; stem cell.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..25 /label= Sig\_peptide  
 FT Protein 26..128  
 FT /label= Mat\_protein  
 XX  
 PN US5643770-A.  
 PD 01-JUL-1997.  
 XX  
 PF 21-JUL-1994; 94US-0278630.  
 XX  
 PR 21-JUL-1994; 94US-0278630.  
 XX  
 PA (ALEX-) ALEXION PHARM INC.  
 XX  
 PI Mason JM, Squinto SP;  
 XX  
 DR WPI: 1997-350243/32.  
 DR N-PSDB; AAT84472.  
 XX  
 PT - for transducing cells in body fluids containing complement  
 PS Disclosure; Column 47-50; 32pp; English.  
 XX  
 CC This protein sequence comprises human CD59, an inhibitor of  
 CC complement C5b-9 protein. Claimed retroviral vector particles  
 CC express a complement inhibitor such as CD59, and are thereby  
 CC protected from inactivation upon exposure to body fluids containing  
 CC complement. Also claimed are: (1) a producer cell producing the  
 CC vector particle; and (2) a chimeric retroviral envelope protein  
 CC (see also AAM26325) with at least part of the N-terminal receptor-  
 CC binding domain removed and replaced by a protein domain having a  
 CC complement inhibitor activity. The vector is used in a claimed  
 CC method for transducing cells in the presence of a body fluid  
 CC containing complement, preferably ex vivo, especially for gene  
 CC therapy, e.g. of hereditary or acquired blood disorders by  
 CC transduction of haematopoietic stem cells.  
 XX  
 SQ Sequence 128 AA;

Query Match 100.0%; Score 96; DB 18; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 9e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 FEHCNFNDVTRRLRENE 17  
 Db 67 fehcnfndvtrrlrene 83

RESULT 16

AAU00688  
 ID AAU00688 standard; Protein: 128 AA.  
 AC AAU00688;  
 XX  
 DT 07-SEP-2001 (first entry)  
 DE Human CD59 protein.  
 XX  
 KW CTLA4; CD59; human; pig; T-cell activation; C5b-9 inhibitory activity;  
 KM C3 inhibitory activity; cellular immune response; xenotransplantation;  
 KW humoral immune response; human serum complement; rodent; mouse; rabbit;  
 KM rat; lagomorph; hare; ungulate; goat; sheep.  
 OS Homo sapiens.  
 XX  
 PN WO200130966-A2.  
 PN  
 PD 03-MAY-2001.  
 XX  
 PF 21-OCT-2000; 2000WO-US29151.  
 XX  
 PR 22-OCT-1999; 99US-0161186.  
 XX  
 PA (ALEX-) ALEXION PHARM INC.  
 XX  
 PI Fodor WL, Pizzolato M;  
 XX  
 DR WPI: 2001-300497/31.  
 DR N-PSDB; AAS00682.  
 XX  
 PT Chimeric protein useful for protecting xenotransplanted tissues by  
 PT C5b-9 and/or C3 inhibitory domain, and a T-cell inhibitory domain -  
 XX  
 PS Example 1; Fig 2E(2); 51pp; English.  
 XX  
 CC The sequence represents human CD59 protein which is fused to human or  
 CC porcine CTLA4 protein to form CTLA4-CD59 chimeric protein. Chimeric  
 CC proteins comprising a domain having C5b-9 and/or C3 inhibitory activity  
 CC (e.g. CTLA4) and a domain having T-cell inhibitory activity (e.g. CD59)  
 CC are capable of inhibiting both cellular immune responses and humoral  
 CC immune responses. These polypeptides and their associated nucleic acids  
 CC are useful for protecting pig cells of tissues and organs from both  
 CC humoral and cellular rejection after xenotransplantation into humans. The  
 CC sequences are capable of conferring resistance to humoral and cellular  
 CC mechanisms of immune attack, to protect against human serum complement  
 CC and to inhibit T-cell activation. Transgenic animals (for example,  
 CC rodents, e.g. mouse, rat; lagomorphs, e.g. rabbit, hare; and ungulates,  
 CC e.g. pig, goat, sheep) expressing such a chimeric protein on the surfaces  
 CC of their cells would have a higher chance of survival.  
 XX  
 SQ Sequence 128 AA;

Query Match 100.0%; Score 96; DB 22; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 9e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 FEHCNFNDVTRRLRENE 17  
 Db 67 fehcnfndvtrrlrene 83

RESULT 17  
 AAU00685  
 ID AAU00685 standard; Protein: 260 AA.  
 AC AAU00685;  
 XX  
 DT 07-SEP-2001 (first entry)  
 DE Human CTLA4-human CD59 chimeric protein.

```
XX CTLA4; CD59; human; pig; T-cell activation; C5b-9 inhibitory activity;
KM C3 inhibitory activity; cellular immune response; xenotransplantation;
KM humoral immune response; human serum complement; rodent; mouse; rabbit;
KM rat; lagomorph; hare; ungulate; goat; sheep; mutant; mutein.
XX Homo sapiens.
OS Synthetic.
XX MO200130966-A2.
PN 03-MAY-2001.
PD
XX
XX 21-OCT-2000; 2000WO-US29151.
PF
XX 22-OCT-1999; 99US-0161186.
PR
XX (ALEX-) ALEXION PHARM INC.
PA
XX Fodor WL, Pizzolato M;
PI
XX WPI: 2001-300497/31.
DR N-PSDB: AAS00679.
XX
XX Chimeric protein useful for protecting xenotransplanted tissues by
PT inhibiting cellular both humoral and immune responses, comprises a
PT C5b-9 and/or C3 inhibitory domain, and a T-cell inhibitory domain -
XX
XX Example 1; Fig 2B(2); 51pp: English.
XX
XX The sequence represents a CTLA4-CD59 chimeric protein, formed from human
CC CTLA4 protein and human CD59 protein. Chimeric proteins comprising a
CC domain having C5b-9 and/or C3 inhibitory activity (e.g. CTLA4) and a
CC domain having T-cell inhibitory activity (e.g. CD59) are capable of
CC inhibiting both cellular immune responses and humoral immune responses.
CC These polypeptides and their associated nucleic acids are useful for
CC protecting pig cells of tissues and organs from both humoral and cellular
CC rejection after xenotransplantation into humans. The sequences are
CC capable of conferring resistance to humoral and cellular mechanisms of
CC immune attack, to protect against human serum complement and to inhibit
CC T-cell activation. Transgenic animals (for example, rodents, e.g. mouse,
CC rat; lagomorphs, e.g. rabbit, hare; and ungulates, e.g. pig, goat, sheep)
CC expressing such a chimeric protein on the surfaces of their cells would
CC have a higher chance of survival.
CC
XX
XX Sequence 260 AA:
SQ
Query Match 100.0%; Score 96; DB 22; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FEHCNENDVTPRLRENE 17
Db 199 fehcnfndvtrtlrene 215
```

```
RESULT 18
AAU00684
ID AAU00684 standard; Protein: 261 AA.
XX
XX AAU00684;
AC
XX
XX 07-SEP-2001 (first entry)
DT
XX
XX Porcine CTLA4-human CD59 chimeric protein.
DE
XX
XX CTLA4; CD59; human; pig; T-cell activation; C5b-9 inhibitory activity;
KM C3 inhibitory activity; cellular immune response; xenotransplantation;
KM humoral immune response; human serum complement; rodent; mouse; rabbit;
KM rat; lagomorph; hare; ungulate; goat; sheep; mutant; mutein.
XX
XX Chimeric - Homo sapiens.
OS
```

```
OS Chimeric - Sus scrofa.
XX
XX MO200130966-A2.
PN 03-MAY-2001.
PD
XX
XX 21-OCT-2000; 2000WO-US29151.
PF
XX 22-OCT-1999; 99US-0161186.
PR
XX (ALEX-) ALEXION PHARM INC.
PA
XX Fodor WL, Pizzolato M;
PI
XX WPI: 2001-300497/31.
DR N-PSDB: AAS00678.
XX
XX Chimeric protein useful for protecting xenotransplanted tissues by
PT inhibiting cellular both humoral and immune responses, comprises a
PT C5b-9 and/or C3 inhibitory domain, and a T-cell inhibitory domain -
XX
XX Example 3; Fig 2A(2); 51pp: English.
XX
XX The sequence represents a CTLA4-CD59 chimeric protein, formed from
CC porcine CTLA4 protein and human CD59 protein. Chimeric proteins
CC comprising a domain having C5b-9 and/or C3 inhibitory activity (e.g.
CC CTLA4) and a domain having T-cell inhibitory activity (e.g. CD59) are
CC capable of inhibiting both cellular immune responses and humoral immune
CC responses. These polypeptides and their associated nucleic acids are
CC useful for protecting pig cells of tissues and organs from both humoral
CC and cellular rejection after xenotransplantation into humans. The
CC sequences are capable of conferring resistance to humoral and cellular
CC mechanisms of immune attack, to protect against human serum complement
CC and to inhibit T-cell activation. Transgenic animals (for example,
CC rodents, e.g. mouse, rat; lagomorphs, e.g. rabbit, hare; and ungulates,
CC e.g. pig, goat, sheep) expressing such a chimeric protein on the surfaces
CC of their cells would have a higher chance of survival.
CC
XX
XX Sequence 261 AA:
SQ
```

```
Query Match 100.0%; Score 96; DB 22; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FEHCNENDVTPRLRENE 17
Db 200 fehcnfndvtrtlrene 216
```

Search completed: April 2, 2002, 10:44:27  
Job time: 32 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:19:51 ; Search time 93.32 Seconds  
(without alignments)  
50.580 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58

Perfect score: 96  
Sequence: 1 FEHCNFNDVYTRLENE 17

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 50 summaries

Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep:\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep:\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep:\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep:\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep:\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep:\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep:\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep:\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep:\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep:\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep:\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep:\*  
24: /cgn2\_6/ptodata/2/paa/US060\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	70	US-08-952-333C-28	Sequence 28, Appl
2	96	100.0	77	US-09-020-393B-3	Sequence 3, Appl
3	96	100.0	80	US-08-043-446A-6	Sequence 6, Appl
4	96	100.0	103	US-07-906-394-3	Sequence 3, Appl
5	96	100.0	103	US-08-021-602-3	Sequence 3, Appl
6	96	100.0	103	US-08-086-549-3	Sequence 3, Appl
7	96	100.0	122	US-09-834-366-13482	Sequence 13482, A
8	96	100.0	122	US-60-197-873-13482	Sequence 13482, A
9	96	100.0	127	US-09-020-393B-1	Sequence 1, Appl

10	96	100.0	128	US-09-834-366-18052	Sequence 18052, A
11	96	100.0	128	US-09-834-366-18056	Sequence 18056, A
12	96	100.0	128	US-09-834-366-18064	Sequence 18064, A
13	96	100.0	128	US-60-197-873-18052	Sequence 18052, A
14	96	100.0	128	US-60-197-873-18056	Sequence 18056, A
15	96	100.0	128	US-60-197-873-18064	Sequence 18064, A
16	89	92.7	78	US-60-160-203-4361	Sequence 4361, Ap
17	89	92.7	78	US-60-169-840-6378	Sequence 6378, Ap
18	72.5	75.5	73	US-60-160-189-7836	Sequence 7836, Ap
19	72.5	75.5	73	US-60-169-867-5425	Sequence 5425, Ap
20	68	70.8	58	PCT-US01-00663-30608	Sequence 30608, A
21	68	70.8	58	US-09-864-761-37475	Sequence 37475, A
22	68	70.8	58	US-60-236-359-19739	Sequence 19739, A
23	68	70.8	75	US-09-020-393B-5	Sequence 5, Appl
24	63	65.6	77	US-09-020-393B-2	Sequence 9833, Ap
25	52	54.2	211	PCT-US01-14827-9833	Sequence 9741, Ap
26	51	53.1	47	US-60-160-189-9741	Sequence 5893, Ap
27	51	53.1	47	US-60-160-203-5893	Sequence 8842, Ap
28	51	53.1	47	US-60-169-840-8842	Sequence 7636, Ap
29	51	53.1	47	US-60-169-867-7696	Sequence 1173, Ap
30	49	51.0	77	US-09-020-393B-6	Sequence 11, Appl
31	49	51.0	77	US-09-020-393B-11	Sequence 2, Appl
32	49	51.0	124	US-09-020-393B-2	Sequence 7, Appl
33	47	49.0	77	US-60-196-174-851	Sequence 851, App
34	46	47.9	50	US-60-196-174-851	Sequence 1174, App
35	46	47.9	50	US-60-234-446-673	Sequence 673, App
36	46	47.9	52	US-60-196-190-1173	Sequence 29464, A
37	46	47.9	57	US-60-196-190-1173	Sequence 4972, Ap
38	46	47.9	294	US-09-252-991A-29464	Sequence 4403, Ap
39	44	45.8	290	US-09-134-001C-4972	Sequence 35, Appl
40	44	45.8	290	US-09-450-969-4403	Sequence 1, Appl
41	44	45.8	15	US-09-188-979-35	Sequence 4548, Ap
42	43	44.8	15	US-09-835-752-1	Sequence 58029, A
43	43	44.8	739	US-60-191-637-5784	Sequence 183630, A
44	43	44.8	739	US-60-191-681-4548	Sequence 21597, A
45	42	43.8	203	US-09-270-767-58029	Sequence 14911, A
46	42	43.8	203	US-09-270-847B-183630	Sequence 14912, A
47	42	43.8	215	US-09-248-796-21597	Sequence 42709, A
48	42	43.8	404	US-09-248-796-14911	
49	42	43.8	415	US-09-248-796-14912	
50	42	43.8	533	US-09-270-767-42709	

#### ALIGNMENTS

RESULT 1

US-08-952-333C-28

Sequence 28, Application US/08952333C

GENERAL INFORMATION:

APPLICANT: Brakenhoff, Rudolf Henrikus; Van Dongen, Augustina

APPLICANT: Antonius Maria Sylvester

TITLE OF INVENTION: Methods for detection and therapy....eptopes

TITLE OF INVENTION: and antibodies

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Frischaut Holte Goodman Langer and Chick

STREET: 767 Third Avenue

City: New York

STATE: New York

COUNTRY: USA

ZIP: 10017

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/952,333C

FILING DATE: 17 FEBRUARY 1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/NL95/00169

FILING DATE: 10 MAY 1995

ATTORNEY/AGENT INFORMATION:  
NAME: Barth, Richard S  
REGISTRATION NUMBER: 28180  
REFERENCE/DOCKET NUMBER: 970686  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 319-4900  
TELEFAX: (212) 319-5101  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 70 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
ORIGINAL SOURCE: protein  
ORGANISM: HCD59 fragment of figure 5  
US-08-952-333C-28

Query Match 100.0%; Score 96; DB 13; Length 70;  
Best Local Similarity 100.0%; Pred. No. 7.2e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTTLRENE 17  
Db 42 FEHCNFDVTTTLRENE 58  
|||||

RESULT 2  
US-09-020-393b-3  
Sequence 3, Application US/09020393b  
GENERAL INFORMATION:  
APPLICANT: Sims, Peter J.  
TITLE OF INVENTION: Compositions and Methods to Inhibit the  
NUMBER OF INVENTION: C5b-9 Complex of Complement  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2800 One Atlantic Center, 1201 W. Peachtree  
STREET: St.  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/020,393b  
FILING DATE: 03-FEB-1998  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMR 170  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-873-8794  
TELEFAX: 404-873-8795  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 77 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Human  
US-09-020-393b-3

Query Match 100.0%; Score 96; DB 14; Length 77;  
Best Local Similarity 100.0%; Pred. No. 8e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTTLRENE 17  
Db 42 FEHCNFDVTTTLRENE 58  
|||||

RESULT 3  
US-08-043-446A-6  
Sequence 6, Application US/08043446A  
GENERAL INFORMATION:  
APPLICANT: Boyd, Richard L.  
APPLICANT: Godfrey, Dale I.  
APPLICANT: MacNeil, Ian A.  
APPLICANT: Zlotnik, Albert  
TITLE OF INVENTION: MAMMALIAN DEVELOPMENTAL ANTIGENS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DNAX Research Institute  
STREET: 901 California Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1104

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/043,446A  
FILING DATE: 05-APR-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ching, Edwin P.  
REGISTRATION NUMBER: 34,090  
REFERENCE/DOCKET NUMBER: DX0375  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-852-9196  
TELEFAX: 415-496-1200  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-043-446A-6

Query Match 100.0%; Score 96; DB 4; Length 80;  
Best Local Similarity 100.0%; Pred. No. 8.4e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTTLRENE 17  
Db 51 FEHCNFDVTTTLRENE 67  
|||||

RESULT 4  
US-07-906-394-3  
Sequence 3, Application US/07906394  
GENERAL INFORMATION:  
APPLICANT: Sims, Peter J.  
APPLICANT: Bothwell, Alfred L.M.  
APPLICANT: Elliott, Eileen A.  
APPLICANT: Flavell, Richard A.  
APPLICANT: Madril, Joseph  
APPLICANT: Rollins, Scott  
APPLICANT: Bell, Leonard  
APPLICANT: Squinto, Stephen





REFERENCE/DOCKET NUMBER: OMRF135  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-815-6500  
TELEFAX: 404-815-6555  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 103 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
CLONE: CDS9  
US-08-086-549-3

Query Match  
Best Local Similarity 100.0%; Score 96; DB 4; Length 103;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNENDVTTRLRENE 17  
|||||  
DB 42 FEHCNENDVTTRLRENE 58

RESULT 7  
US-09-834-366-13482  
Sequence 13482, Application US/09834366  
GENERAL INFORMATION:  
APPLICANT: Bejani, Stephanie  
APPLICANT: Tanaka, Hiroaki  
APPLICANT: Dumas Milne Edwards, Jean Baptiste  
APPLICANT: Jobert, Severin  
APPLICANT: Giordano, Jean-Yves  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: 81.US2.REG  
CURRENT APPLICATION NUMBER: US/09/834,366  
PRIOR FILING DATE: 2001-04-13  
PRIOR APPLICATION NUMBER: US 60/197,873  
NUMBER OF SEQ ID NOS: 2000-04-18  
SOFTWARE: Patent.pm  
SEQ ID NO 13482  
LENGTH: 122  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: -20...-1  
US-09-834-366-13482

Query Match  
Best Local Similarity 100.0%; Score 96; DB 22; Length 122;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNENDVTTRLRENE 17  
|||||  
DB 67 FEHCNENDVTTRLRENE 83

RESULT 8  
US-60-197-873-13482  
Sequence 13482, Application US/60197873  
GENERAL INFORMATION:  
APPLICANT: Bejani, Stephanie  
APPLICANT: Tanaka, Hiroaki  
APPLICANT: Dumas Milne Edwards, Jean Baptiste

APPLICANT: Jobert, Severin  
APPLICANT: Giordano, Jean-Yves  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: 81.US1.PRO  
CURRENT APPLICATION NUMBER: US/60/197,873  
CURRENT FILING DATE: 2000-04-18  
NUMBER OF SEQ ID NOS: 52153  
SOFTWARE: Patent.pm  
SEQ ID NO 13482  
LENGTH: 122  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: -20...-1  
US-60-197-873-13482

Query Match  
Best Local Similarity 100.0%; Score 96; DB 24; Length 122;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNENDVTTRLRENE 17  
|||||  
DB 67 FEHCNENDVTTRLRENE 83

RESULT 9  
US-09-020-393b-1  
Sequence 1, Application US/09020393b  
GENERAL INFORMATION:  
APPLICANT: Sims, Peter J.  
TITLE OF INVENTION: Compositions and Methods to Inhibit the  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2800 One Atlantic Center, 1201 W. Peachtree  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/020,393b  
FILING DATE: 03-FEB-1998  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRF 170  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-873-8794  
TELEFAX: 404-873-8795  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 127 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Human  
US-09-020-393b-1

Query Match 100.0%; Score 96; DB 14; Length 127;

		Best Local Similarity	100.0%;	Pred. No.	1.4e-07;						
		Matches	17;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0.
Qy	1	FEHCNFNDVYTRLRNE	17								
Db	66	FEHCNFNDVYTRLRNE	82								

```

US-09-834-366-18052
US-09-834-366-18052
10
US-09-834-366-18052
Sequence 18052, Application US/09834366
GENERAL INFORMATION:
APPLICANT: Benjamin, Stephane
APPLICANT: Tanaka, Hiroaki
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Jobert, Severin
APPLICANT: Giordano, Jean-Yves
TITLE OF INVENTION: ESTs and Encoded Human Proteins
FILE REFERENCE: 81.052.REG
CURRENT APPLICATION NUMBER: US/09/834,366
CURRENT FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/197,873
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 52153
SOFTWARE: Patent.pm
SEQ ID NO 18052
LENGTH: 128
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -20..-1
US-09-834-366-18052

```

Query Match	100.0%	Score 96;	DB 22;	Length 128;
Best Local Similarity	100.0%	Pred. No. 1,4e-07;		
Matches 17; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	FEHCNPNNDYTRLRENE	17	
DB	67	FEHCNPNNDYTRLRENE	83	

```

RESULT 11
US-09-834-366-18056
; Sequence 18056, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US2.REG
; CURRENT APPLICATION NUMBER: US/09/834,366
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 18056
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -20...-1
; US-09-834-366-18056

```

Query Match	100.0%;	Score 96;	DB 22;	Length 128;
Best Local Similarity	100.0%;	Pred. No. 1.4e-07;		

Matches	17	Conservative	0	Mismatches	0	Indels	0	Gaps	0
Qy	1	FEHCNFNDVYTRLRENE	17						
Db	67	FEHCNFNDVYTRLRENE	83						

```

US-09      12
US-09-834-366-18064
; Sequence 18064, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Bejani, Stephanie
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Mline Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins
; FILE REFERENCE: 81.US2.REG
; CURRENT APPLICATION NUMBER: US/09/834,366
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 18064
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -20..-1
US-09-834-366-18064

```

Query Match	100.0%	Score 96	DB 22	Length 128
Best Local Similarity	100.0%	Pred. No. 14e-07		
Matches 17	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	FEHCNFDVYTR	RENE 17	
Db	67	FEHCNFDVYTR	RENE 83	

```

RESULT 13
US-60-197-873-18052
; Sequence 18052, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejantin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dunas Milne Edwards, Jean Baptiste
; APPLICANT: Joberit, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins
; FILE REFERENCE: 81.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/197,873
; CURRENT FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 18052
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: ~20...1
US-60-197-873-18052

```

Query Match	100.0%	Score 96;	DB 24;	Length 128;
Best Local Similarity	100.0%	Pred. No. 1,4e-07;		
Matches 17; Conservative	0;	Mismatches	0;	Indels 0;
QY	1	FEHCNFNDVTTPLRENE	17	

Db

67 FEHCNFDVTTLRRENE 83

## RESULT 14

```
/ Sequence 18056, Application US/60197873
/ GENERAL INFORMATION:
/ APPLICANT: Bejani, Stephane
/ APPLICANT: Tanaka, Hiroaki
/ APPLICANT: Dumas Milne Edwards, Jean Baptiste
/ APPLICANT: Jobert, Severin
/ TITLE OF INVENTION: ESTs and Encoded Human Proteins.
/ FILE REFERENCE: 81.US1.PRO
/ CURRENT APPLICATION NUMBER: US/60/197,873
/ NUMBER OF SEQ ID NOS: 52153
/ SOFTWARE: Patent.pm
/ SEQ ID NO 18056
/ LENGTH: 128
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SIGNAL
/ LOCATION: -20...-1
US-60-197-873-18056
```

## Query Match

Best Local Similarity 100.0%; Score 96; DB 24; Length 128;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTLRRENE 17  
Db 67 FEHCNFDVTTLRRENE 83

## RESULT 15

```
/ Sequence 18064, Application US/60197873
/ GENERAL INFORMATION:
/ APPLICANT: Bejani, Stephane
/ APPLICANT: Tanaka, Hiroaki
/ APPLICANT: Dumas Milne Edwards, Jean Baptiste
/ APPLICANT: Jobert, Severin
/ TITLE OF INVENTION: ESTs and Encoded Human Proteins.
/ FILE REFERENCE: 81.US1.PRO
/ CURRENT APPLICATION NUMBER: US/60/197,873
/ NUMBER OF SEQ ID NOS: 52153
/ SOFTWARE: Patent.pm
/ SEQ ID NO 18064
/ LENGTH: 128
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SIGNAL
/ LOCATION: -20...-1
US-60-197-873-18064
```

## Query Match

Best Local Similarity 100.0%; Score 96; DB 24; Length 128;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTLRRENE 17  
Db 67 FEHCNFDVTTLRRENE 83

## RESULT 16

US-60-160-203-4361

```
/ Sequence 4361, Application US/60160203
/ GENERAL INFORMATION:
/ APPLICANT: BONAZZI, VIVIAN
/ TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
/ FILE REFERENCE: CL000116
/ CURRENT APPLICATION NUMBER: US/60/160,203
/ NUMBER OF SEQ ID NOS: 6374
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4361
/ LENGTH: 78
/ TYPE: PRT
/ ORGANISM: HUMAN
/ FEATURE:
/ NAME/KEY: VARIANT
/ LOCATION: (1)...(78)
/ OTHER INFORMATION: Xaa = Any Amino Acid
US-60-160-203-4361
```

## Query Match

Best Local Similarity 92.7%; Score 89; DB 24; Length 78;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTLRRENE 17  
Db 17 FEHCNFDVTTLRRENE 33

## RESULT 17

```
/ Sequence 6378, Application US/60169840
/ GENERAL INFORMATION:
/ APPLICANT: Bonazzi, Vivien
/ TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
/ FILE REFERENCE: CL000164
/ CURRENT APPLICATION NUMBER: US/60/169,840
/ NUMBER OF SEQ ID NOS: 9628
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 6378
/ LENGTH: 78
/ TYPE: PRT
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: VARIANT
/ LOCATION: (1)...(78)
/ OTHER INFORMATION: Xaa = Any Amino Acid
US-60-169-840-6378
```

## Query Match

Best Local Similarity 92.7%; Score 89; DB 24; Length 78;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTLRRENE 17  
Db 17 FEHCNFDVTTLRRENE 33

## RESULT 18

```
/ Sequence 7836, Application US/60160189
/ GENERAL INFORMATION:
/ APPLICANT: BONAZZI, VIVIAN
/ TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS
/ TITLE OF INVENTION: AND USES THEREOF
US-60-160-189-7836
```

```
; FILE REFERENCE: CL0000112
; CURRENT APPLICATION NUMBER: US/60/160,189
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 10162
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7836
; LENGTH: 73
; TYPE: PRF
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(73)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-160-189-7836

Query Match          75.5%; Score 72.5; DB 24; Length 73;
Best Local Similarity 83.3%; Pred. No. 0.00056;
Matches 15; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Oy 1 FEHCNF-NDVTTRLRENE 17
    ||||| |||||
Db 11 FEHCNFRXYVTRLRENE 28

RESULT 19
US-60-169-867-5425
; Sequence 5425, Application US/60169867
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
; FILE REFERENCE: C1000160
; CURRENT APPLICATION NUMBER: US/60/169,867
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 8230
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5425
; LENGTH: 73
; TYPE: PRF
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(73)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-169-867-5425

Query Match          75.5%; Score 72.5; DB 24; Length 73;
Best Local Similarity 83.3%; Pred. No. 0.00056;
Matches 15; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Oy 1 FEHCNF-NDVTTRLRENE 17
    ||||| |||||
Db 11 FEHCNFRXYVTRLRENE 28

RESULT 20
PCT-US01-00663-30608
; Sequence 30608, Application PC/TUS0100663
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 7
; CURRENT APPLICATION NUMBER: PCT/US01/00663
; CURRENT FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/180,312
```

```
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38837
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 30608
; LENGTH: 58
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL049629.14
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 16
; OTHER INFORMATION: EST HUMAN HIT: A0139072.1, EVALU6 7.00e-29
; OTHER INFORMATION: SWISSPROT HIT: P13987, EVALU6 6.00e-30
PCT-US01-00663-30608

Query Match          70.8%; Score 68; DB 1; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 NFNDVTTRLRENE 17
    ||||| |||||
Db 1 NFNDVTTRLRENE 13

RESULT 21
US-09-864-761-37475
; Sequence 37475, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
```

PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 37475  
LENGTH: 58  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL049629.14  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.6  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.6  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.2  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.7  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 16  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.6  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.9  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.9  
OTHER INFORMATION: EST\_HUMAN HIT: AU139072.1, EVALUE 7.00e-29  
OTHER INFORMATION: SWISSPROT HIT: P13987, EVALUE 6.00e-30  
US-09-864-761-37475

Query Match  
Best Local Similarity 70.8%; Score 68; DB 22; Length 58;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 NFNDVTTLRRENE 17  
|||||  
Db 1 NFNDVTTLRRENE 13

RESULT 22  
US-60-236-359-19739  
Sequence 19739, Application US/60236359  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEOTIC ACID PROBES USEFUL FOR  
FILE REFERENCE: MDHMR-4P  
CURRENT FILING DATE: 2000-09-27  
CURRENT APPLICATION NUMBER: US/60/236,359  
PRIOR FILING DATE: 2000-05-26  
NUMBER OF SEQ ID NOS: 21709  
SOFTWARE: Molecular Dynamics Sequence Listing Engine  
SEQ ID NO 19739  
LENGTH: 58  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL049629.14  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.6  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.6  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.2  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.7  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 16

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.6  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.9  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.9  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.6  
OTHER INFORMATION: EST\_HUMAN HIT: AW374176.1, EVALUE 4.00e-29  
US-60-236-359-19739  
SWISSPROT HIT: P13987, EVALUE 6.00e-30

Query Match  
Best Local Similarity 70.8%; Score 68; DB 24; Length 58;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 NFNDVTTLRRENE 17  
|||||  
Db 1 NFNDVTTLRRENE 13

RESULT 23  
US-09-020-393B-4  
Sequence 4, Application US/09020393B  
GENERAL INFORMATION:  
APPLICANT: Sims, Peter J.  
TITLE OF INVENTION: Compositions and Methods to Inhibit the  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Padst  
STREET: 2800 One Atlantic Center, 1201 W. Peachtree  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 03-FEB-1998  
APPLICATION NUMBER: US/09/020,393B  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Padst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OARF 170  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-873-8794  
TELEFAX: 404-873-8795  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 75 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Peptide  
HYPOTHEICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Badoon  
US-09-020-393B-4

Query Match  
Best Local Similarity 70.8%; Score 68; DB 14; Length 75;  
Matches 11; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Oy 1 FENNDVTTLRRENE 17  
|:|||||:|  
Db 42 FANFNDDSTLTKENE 58

RESULT 24

US-09-020-393B-5  
; Sequence 5, Application US/09020393B  
; GENERAL INFORMATION:  
; APPLICANT: Sims, Peter J.  
; TITLE OF INVENTION: Compositions and Methods to Inhibit the  
; TITLE OF INVENTION: C5b-9 Complex of Complement  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patrea L. Pabst  
; STREET: 2800 One Atlantic Center, 1201 W. Peachtree  
; STREET: St.  
; CITY: Atlanta  
; STATE: GA  
; COUNTRY: USA  
; ZIP: 30309-3450  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/020,393B  
; FILING DATE: 03-FEB-1998  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: OMR# 170  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404-873-8794  
; TELEFAX: 404-873-8795  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 77 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: African green monkey  
; US-09-020-393B-5

Query Match 65.6%; Score 63; DB 14; Length 77;  
Best Local Similarity 58.8%; Pred. No. 0.022;  
Matches 10; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 FEHCNENDVTTRLRENE 17  
+ :|||||:| :| :|  
Db 42 FANCFNDISTLLKESE 58

RESULT 25  
PCT-US01-14827-9833  
; Sequence 9833, Application PC/TUS0114827  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 21272-104  
; CURRENT APPLICATION NUMBER: PCT/US01/14827  
; CURRENT FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: 09/577,408  
; PRIOR FILING DATE: 2000-05-18  
; NUMBER OF SEQ ID NOS: 16102  
; SOFTWARE: Custom  
; SEQ ID NO 9833  
; LENGTH: 211  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (43)..(52)

; OTHER INFORMATION: Ly-6 / u-PAR domain proteins domain identified by eMATRIX.  
; OTHER INFORMATION: accession number BL00983A, p-value=7.480e-09, raw score of 5.8  
; NAME/KEY: DOMAIN  
; LOCATION: (32)..(160)  
; OTHER INFORMATION: u-PAR/Ly-6 domain identified by Pnam, accession name  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(211)  
; OTHER INFORMATION: Xaa = X or \* as defined in Table 2  
; PCT-US01-14827-9833

Query Match 54.2%; Score 52; DB 1; Length 211;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHCNEND 8  
+ |||||  
Db 103 FEHCNEND 110

Search completed: April 2, 2002, 09:22:42  
Job time: 171 sec

\_\_\_\_\_



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:19:51 ; Search time 16.47 Seconds  
(without alignments)  
73.774 Million cell updates/sec

```
Title: US-09-020-393B-3_COPY_42_58
Perfect score: 96
Sequence: 1 FEHCNFDVTRLRENE 17
```

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 348507 seqs, 71473746 residues

Total number of hits satisfying chosen parameters: 348507

Minimum	DB	seq	length: 0
Maximum	DB	seq	length: 2000000000

Post-processing: Minimum Match 0%

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 50 summaries
```

```
Database : Pending_Patents_AA_NEM:*
1: /cgn2_6/ptodata/1/paa/PCn_NEM_COMB pep: *
2: /cgn2_6/ptodata/1/paa/US06_NEM_COMB pep: *
3: /cgn2_6/ptodata/1/paa/US07_NEM_COMB pep: *
4: /cgn2_6/ptodata/1/paa/US08_NEM_COMB pep: *
5: /cgn2_6/ptodata/1/paa/US09_NEM_COMB pep: *
6: /cgn2_6/ptodata/1/paa/US09_NEM_COMB pep1: *
7: /cgn2_6/ptodata/1/paa/US10_NEM_COMB pep: *
8: /cgn2_6/ptodata/1/paa/US10_NEM_COMB pep: *
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	96	100.0	70	6	US-09-612-114-39	Sequence 42, Appl
2	96	100.0	71	6	US-09-612-114-42	Sequence 49, Appl
3	96	100.0	77	6	US-09-612-114-37	Sequence 37, Appl
4	96	100.0	82	6	US-09-612-114-40	Sequence 40, Appl
5	96	100.0	83	6	US-09-612-114-41	Sequence 41, Appl
6	96	100.0	88	6	US-09-612-114-43	Sequence 45, Appl
7	96	100.0	99	6	US-09-612-114-45	Sequence 43, Appl
8	96	100.0	100	6	US-09-612-114-44	Sequence 44, Appl
9	96	100.0	128	6	US-09-612-114-44	Sequence 20, Appl
10	96	100.0	128	6	US-09-614-150-5766	Sequence 5766, Appl
11	41.5	43.2	236	5	US-09-991-630-12	Sequence 12, Appl
12	41.5	43.2	236	5	US-09-991-630-15	Sequence 15, Appl
13	41.5	43.2	236	6	US-09-905-129-12	Sequence 12, Appl
14	41.5	43.2	236	6	US-09-905-129-15	Sequence 15, Appl
15	40	41.7	220	6	US-09-708-427-16838	Sequence 16838, A
16	40	41.7	244	6	US-09-897-516-6395	Sequence 6395, A
17	40	41.7	244	6	US-09-708-427-16837	Sequence 16837, A
18	40	41.7	244	8	US-60-336-049-36	Sequence 36, Appl
19	40	41.7	262	7	US-10-092-302-711	Sequence 711, Appl
20	40	41.7	262	7	US-10-092-302-951	Sequence 951, Appl
21	40	41.7	272	6	US-09-708-427-56000	Sequence 56000, A
22	40	41.7	299	6	US-09-708-427-55999	Sequence 55999, A
23	40	41.7	394	1	PCT-US02-03987-15325	Sequence 15325, A
24	40	41.7	394	7	US-10-032-585-7381	Sequence 7381, A
25	40	41.7	394	7	US-10-072-851-15325	Sequence 15325, A

## ALIGNMENTS

26	40	41.7	403	6	US-09-708-427-55395	Sequence 5395, Ap
27	40	41.7	429	6	US-09-708-427-55394	Sequence 5394, Ap
28	40	41.7	497	6	US-09-708-427-55393	Sequence 5393, Ap
29	40	41.7	741	6	US-09-614-150-40656	Sequence 4056, Ap
30	40	41.7	901	6	US-09-614-150-2010	Sequence 2010, Ap
31	39	40.6	121	6	US-09-897-515-5831	Sequence 5831, Ap
32	39	40.6	123	6	US-09-673-033-1	Sequence 1, Appl
33	39	40.6	135	6	US-09-897-515-4600	Sequence 4600, Ap
34	39	40.6	191	8	US-60-361-742-25	Sequence 25, Appl
35	39	40.6	210	5	US-09-302-626B-169	Sequence 169, Appl
36	39	40.6	213	6	US-09-897-515-7509	Sequence 7509, Ap
37	39	40.6	222	5	US-09-302-626B-174	Sequence 174, Appl
38	39	40.6	240	5	US-09-302-626B-170	Sequence 170, Appl
39	39	40.6	260	5	US-09-589-927-12	Sequence 12, Appl
40	39	40.6	260	5	US-09-589-987-12	Sequence 12, Appl
41	39	40.6	260	5	US-09-618-301-12	Sequence 12, Appl
42	39	40.6	260	5	US-09-618-519-12	Sequence 12, Appl
43	39	40.6	268	7	US-10-029-386-33039	Sequence 33039, A
44	39	40.6	374	6	US-09-897-515-7828	Sequence 7828, Ap
45	39	40.6	380	6	US-09-760-446A-1772	Sequence 1772, Ap
46	39	40.6	571	5	US-09-302-626B-26	Sequence 26, Appl
47	39	40.6	672	7	US-10-010-920-31	Sequence 31, Appl
48	39	40.6	672	8	US-60-338-287-31	Sequence 31, Appl
49	39	40.6	690	5	US-09-302-626B-171	Sequence 171, Appl
50	39	40.6	701	6	US-09-614-150-2730	Sequence 2730, Ap

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	100.0%	Score 96:	DB 6:	Length 70:
Best Local Similarity	100.0%	Pred. No.	3e-09:	-
Matches 17; Conservative	0;	Mismatches	0;	Indels 0
				Gaps 0

```

Db 42 FEHCNENDVTRLRENE 58

```

RESULT 2  
US-09-612-314-42  
; Sequence 42, Application US/09612314  
; GENERAL INFORMATION:  
; APPLICANT: SMITH, Richard, Anthony, Godwin

```

; APPLICANT: DODD, Ian
; TITLE OF INVENTION: MOSSAKOWSKA, Danuta, Ewa, Irena
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; FILE REFERENCE: 37945-0004
; CURRENT APPLICATION NUMBER: US/09/612,314
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Protein APT2061
US-09-612-314-42

```

```

Query Match
Best Local Similarity 100.0%; Score 96; DB 6; Length 71;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FEHCNENDVYTRLENE 17
Db 42 FEHCNENDVYTRLENE 58

```

```

RESULT 3
US-09-612-314-37
; Sequence 37, Application US/09612314
; GENERAL INFORMATION:
; APPLICANT: SMITH, Richard, Anthony, Godwin
; APPLICANT: DODD, Ian
; TITLE OF INVENTION: MOSSAKOWSKA, Danuta, Ewa, Irena
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; FILE REFERENCE: 37945-0004
; CURRENT APPLICATION NUMBER: US/09/612,314
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Protein APT631
US-09-612-314-37

```

```

Query Match
Best Local Similarity 100.0%; Score 96; DB 6; Length 77;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FEHCNENDVYTRLENE 17
Db 42 FEHCNENDVYTRLENE 58

```

```

RESULT 4
US-09-612-314-40

```

```

; Sequence 40, Application US/09612314
; GENERAL INFORMATION:
; APPLICANT: SMITH, Richard, Anthony, Godwin
; APPLICANT: DODD, Ian
; TITLE OF INVENTION: MOSSAKOWSKA, Danuta, Ewa, Irena
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; FILE REFERENCE: 37945-0004
; CURRENT APPLICATION NUMBER: US/09/612,314
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Protein APT2060
US-09-612-314-40

```

```

Query Match
Best Local Similarity 100.0%; Score 96; DB 6; Length 82;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FEHCNENDVYTRLENE 17
Db 42 FEHCNENDVYTRLENE 58

```

```

RESULT 5
US-09-612-314-41
; Sequence 41, Application US/09612314
; GENERAL INFORMATION:
; APPLICANT: SMITH, Richard, Anthony, Godwin
; APPLICANT: DODD, Ian
; TITLE OF INVENTION: MOSSAKOWSKA, Danuta, Ewa, Irena
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; FILE REFERENCE: 37945-0004
; CURRENT APPLICATION NUMBER: US/09/612,314
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Protein APT635
US-09-612-314-41

```

```

Query Match
Best Local Similarity 100.0%; Score 96; DB 6; Length 83;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FEHCNENDVYTRLENE 17
Db 43 FEHCNENDVYTRLENE 59

```

```

RESULT      6
US-09-612-314-45
: Sequence 45, Application US/09612314
: GENERAL INFORMATION:
: APPLICANT: SMITH, Richard, Anthony, Godwin
: APPLICANT: DODD, Ian
: APPLICANT: MOSSAKOWSKA, Danuta, Ewa, Irena
: TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
: FILE REFERENCE: 37945-0004
: CURRENT APPLICATION NUMBER: US/09/612.314
: PRIOR FILING DATE: 2000-07-07
: PRIOR APPLICATION NUMBER: US 09/214,913
: PRIOR FILING DATE: 1999-03-16
: PRIOR APPLICATION NUMBER: PCT/EP97/03715
: PRIOR FILING DATE: 1997-07-08
: PRIOR APPLICATION NUMBER: GB 96 148 71.3
: PRIOR FILING DATE: 1996-07-15
: NUMBER OF SEQ ID NOS: 67
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 45
: LENGTH: 88
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (71)..(72)
: OTHER INFORMATION: Residues 1-71 are a first polypeptide chain and
: OTHER INFORMATION: Residues 72-88 are a second polypeptide chain
: OTHER INFORMATION: linked by a disulphide bond formed between the
: NAME/KEY: UNSURE
: LOCATION: (88)
: OTHER INFORMATION: The C-terminal glycine is linked to an
: OTHER INFORMATION: NH-[Myristoyl] group
: OTHER INFORMATION: Description of Artificial Sequence: Protein APT2065
US-09-612-314-45

Query Match      100.0%; Score 96; DB 6; Length 88;
Best Local Similarity 100.0%; Pred. NO. 3.9e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY      1 FEHCNFNDVTRIRENE 17
        |||
Db      42 FEHCNFNDVTRIRENE 58

RESULT      7
US-09-612-314-43
: Sequence 43, Application US/09612314
: GENERAL INFORMATION:
: APPLICANT: SMITH, Richard, Anthony, Godwin
: APPLICANT: DODD, Ian
: APPLICANT: MOSSAKOWSKA, Danuta, Ewa, Irena
: TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
: FILE REFERENCE: 37945-0004
: CURRENT APPLICATION NUMBER: US/09/612.314
: CURRENT FILING DATE: 2000-07-07
: PRIOR APPLICATION NUMBER: US 09/214,913
: PRIOR FILING DATE: 1999-03-16
: PRIOR APPLICATION NUMBER: PCT/EP97/03715
: PRIOR FILING DATE: 1997-07-08
: PRIOR APPLICATION NUMBER: GB 96 148 71.3
: PRIOR FILING DATE: 1996-07-15
: NUMBER OF SEQ ID NOS: 67
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 43
: LENGTH: 99
: TYPE: PRT
: ORGANISM: Artificial Sequence

```

```

: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (82)..(83)
: OTHER INFORMATION: Residues 1-82 are a first polypeptide chain and
: OTHER INFORMATION: residues 83-99 are a second polypeptide chain
: OTHER INFORMATION: linked by a disulphide bond formed between the
: OTHER INFORMATION: cysteines at positions 82 and 83
: NAME/KEY: UNSURE
: LOCATION: (99)
: OTHER INFORMATION: The C-terminal glycine is linked to an
: OTHER INFORMATION: NH-[Wyrstoyl] group
: OTHER INFORMATION: Description of Artificial Sequence: Protein APT2062
US-09-612-314-43

Query Match          100.0%; Score 96; DB 6; Length 99;
Best Local Similarity 100.0%; Pred. No. 4,4e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FEHCNFNDVTRRLRENE 17
        |||
Db      42 FEHCNFNDVTRRLRENE 58

RESULT      8
US-09-612-314-44
: Sequence 44, Application US/09612314
: GENERAL INFORMATION:
: APPLICANT: SMITH, Richard, Anthony, Godwin
: APPLICANT: DODD, Ian
: APPLICANT: MOSSAKOMSKA, Danuta, Ewa, Irena
: TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
: FILE REFERENCE: 3/945-0004
: CURRENT APPLICATION NUMBER: US/09/612,314
: CURRENT FILING DATE: 2000-07-07
: PRIOR APPLICATION NUMBER: US 09/214,913
: PRIOR FILING DATE: 1999-03-16
: PRIOR APPLICATION NUMBER: PCT/EP97/03715
: PRIOR FILING DATE: 1997-07-08
: PRIOR APPLICATION NUMBER: GB 96 148 71.3
: PRIOR FILING DATE: 1996-07-15
: NUMBER OF SEQ ID NOS: 67
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 44
: LENGTH: 100
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (83)..(84)
: OTHER INFORMATION: Residues 1-83 are a first polypeptide chain and
: OTHER INFORMATION: residues 84-100 are a second polypeptide chain
: OTHER INFORMATION: linked by a disulphide bond formed between the
: OTHER INFORMATION: cysteines at positions 83 and 84
: NAME/KEY: UNSURE
: LOCATION: (100)
: OTHER INFORMATION: The C-terminal glycine is linked to an
: OTHER INFORMATION: NH-[Wyrstoyl] group
: OTHER INFORMATION: Description of Artificial Sequence: Protein APT2063
US-09-612-314-44

Query Match          100.0%; Score 96; DB 6; Length 100;
Best Local Similarity 100.0%; Pred. No. 4,5e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FEHCNFNDVTRRLRENE 17
        |||
Db      43 FEHCNFNDVTRRLRENE 59

RESULT      9

```

```
US-09-673-032-20
; Sequence 20, Application US/09673032
; GENERAL INFORMATION:
; APPLICANT: MORGAN, BRYAN P.
; APPLICANT: RUSHMERE, NEIL K.
; APPLICANT: HINCHLIFE, STEWART J.
; APPLICANT: VAN DEN BERG, CARMEN W.
; TITLE OF INVENTION: MODIFIED BIOLOGICAL MATERIAL
; FILE REFERENCE: WN/KH/JT/MCM
; CURRENT FILING DATE: 2000-12-06
; PRIOR FILING DATE: 1999-04-08
; PRIOR FILING DATE: 1998-04-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-673-032-20
```

```
Query Match
Best Local Similarity 100.0%; Score 96; DB 6; Length 128;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTTLRENE 17
Db 67 FEHCNFDVTTTLRENE 83
```

```
RESULT 10
US-09-614-150-5766
; Sequence 5766, Application US/09614150
; GENERAL INFORMATION:
; APPLICANT: Ventef, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/09/614,150
; CURRENT FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5766
; LENGTH: 739
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-09-614-150-5766
```

```
Query Match
Best Local Similarity 44.8%; Score 43; DB 6; Length 739;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 2 EHCNFDVTTTLREN 16
Db 374 EHCNFDVTTTLREN 388
```

```
RESULT 11
US-09-991-630-12
; Sequence 12, Application US/09991630
; GENERAL INFORMATION:
; APPLICANT: Elnat, et al
; TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODUCTS T
; FILE REFERENCE: 540579-2007.3
; CURRENT APPLICATION NUMBER: US/09/991,630
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 09/905,129
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 09/802,318
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 09/729,485
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(236)
; OTHER INFORMATION: 'x' can be any amino acid
US-09-991-630-12
```

```
Query Match
Best Local Similarity 43.2%; Score 41.5; DB 5; Length 236;
Matches 10; Conservative 2; Mismatches 3; Indels 19; Gaps 1;

QY 3 HCNFDVTTTLRENE 17
Db 42 HCNFDVTTTLRENE 75
```

```
RESULT 12
US-09-991-630-15
; Sequence 15, Application US/09991630
; GENERAL INFORMATION:
; APPLICANT: Elnat, et al
; TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODUCTS TH
; FILE REFERENCE: 540579-2007.3
; CURRENT APPLICATION NUMBER: US/09/991,630
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 09/905,129
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 09/802,318
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 09/729,485
; PRIOR FILING DATE: 2000-12-04
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(236)
; OTHER INFORMATION: 'x' can be any amino acid
US-09-991-630-15
```

```
Query Match          43.2%; Score 41.5; DB 5; Length 236;
Best Local Similarity 29.4%; Pred. No. 20;
Matches 10; Conservative 2; Mismatches 3; Indels 19; Gaps 1;

QY 3 HCNFNDVT-----TRLRENE 17
|||:|:|
Db 42 HCTFRDLTSPDGPANVERVNLGYNLSLTRLTEND 75

RESULT 13
US-09-905-129-12
; Sequence 12, Application US/09905129
; GENERAL INFORMATION:
; APPLICANT: Elnat, et al
; TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODUCTS THEREOF
; FILE REFERENCE: 540579-2007.2
; CURRENT APPLICATION NUMBER: US/09/905,129
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 09/802,318
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/207,821
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 60/084,944
; PRIOR FILING DATE: 1998-05-11
; PRIOR APPLICATION NUMBER: 60/085,673
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 12
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(236)
; OTHER INFORMATION: 'x' can be any amino acid
US-09-905-129-12

Query Match          43.2%; Score 41.5; DB 6; Length 236;
Best Local Similarity 29.4%; Pred. No. 20;
Matches 10; Conservative 2; Mismatches 3; Indels 19; Gaps 1;

QY 3 HCNFNDVT-----TRLRENE 17
|||:|:|
Db 42 HCTFRDLTSPDGPANVERVNLGYNLSLTRLTEND 75

RESULT 14
US-09-905-129-15
; Sequence 15, Application US/09905129
; GENERAL INFORMATION:
; APPLICANT: Elnat, et al
; TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODUCTS THEREOF
; FILE REFERENCE: 540579-2007.2
; CURRENT APPLICATION NUMBER: US/09/905,129
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 09/802,318
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/207,821
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 60/084,944
; PRIOR FILING DATE: 1998-05-11
; PRIOR APPLICATION NUMBER: 60/085,673
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 15
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Mus musculus

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(236)
; OTHER INFORMATION: 'x' can be any amino acid
US-09-905-129-15

Query Match          43.2%; Score 41.5; DB 6; Length 236;
Best Local Similarity 29.4%; Pred. No. 20;
Matches 10; Conservative 2; Mismatches 3; Indels 19; Gaps 1;

QY 3 HCNFNDVT-----TRLRENE 17
|||:|:|
Db 42 HCTFRDLTSPDGPANVERVNLGYNLSLTRLTEND 75

RESULT 15
US-09-708-427-16838
; Sequence 16838, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 16838
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..220
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..220
; OTHER INFORMATION: Ceres Seq. ID 1831334
US-09-708-427-16838

Query Match          41.7%; Score 40; DB 6; Length 220;
Best Local Similarity 46.2%; Pred. No. 34;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTRL 13
|:||||:|:|
Db 60 FKHCNFSFVROL 72

RESULT 16
US-09-897-516-6395
; Sequence 6395, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvaiz, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spirdonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 6395
; LENGTH: 240
; TYPE: PRT
```

; ORGANISM: Xenorhabdus sp.  
US-09-897-516-6395

Query Match  
Best Local Similarity 41.7%; Score 40; DB 6; Length 240;  
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 FHCNFDVTRLRENE 17  
| | | | | : |  
Db 62 HDTFNDVLRNLPHE 76

## RESULT 17

US-09-708-427-16837  
; Sequence 16837, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16837  
; LENGTH: 244  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..244  
; OTHER INFORMATION: Xaa is any amino acid  
; LOCATION: 1..244  
; OTHER INFORMATION: Ceres Seq. ID 1831333  
US-09-708-427-16837

## Query Match

Best Local Similarity 41.7%; Score 40; DB 6; Length 244;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 FHCNFDVTRL 13  
| | | | | : |  
Db 84 FKHCNFSFVROL 96

## RESULT 18

US-60-336-049-36  
; Sequence 36, Application US/60336049  
; GENERAL INFORMATION:  
; APPLICANT: Adam, Luc  
; APPLICANT: Creelman, Robert  
; APPLICANT: Dubell, Arnold  
; APPLICANT: Heard, Jacqueline  
; APPLICANT: Jiang, Cai-Zhong  
; APPLICANT: Keddie, James  
; APPLICANT: Pilgrim, Marsha  
; APPLICANT: Ratcliffe, Oliver  
; APPLICANT: Reuber, Lynne  
; APPLICANT: Riechmann, Jose Luis  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: Genes for Modifying Plant Traits VI  
; FILE REFERENCE: M81-0037  
; CURRENT APPLICATION NUMBER: US/60/336,049  
; CURRENT FILING DATE: 2001-12-05  
; NUMBER OF SEQ ID NOS: 784  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 36  
; LENGTH: 244  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana

US-60-336-049-36

Query Match  
Best Local Similarity 41.7%; Score 40; DB 8; Length 244;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 FHCNFDVTRL 13  
| | | | | : |  
Db 84 FKHCNFSFVROL 96

## RESULT 19

US-10-092-302-711  
; Sequence 711, Application US/10092302  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT218C1  
; CURRENT APPLICATION NUMBER: US/10/092,302  
; CURRENT FILING DATE: 2002-03-07  
; NUMBER OF SEQ ID NOS: 1040  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 711  
; LENGTH: 262  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-302-711

## Query Match

Best Local Similarity 41.7%; Score 40; DB 7; Length 262;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FHCNFDVTRL 11  
| | | | | : |  
Db 51 FEECYFEDVTS 61

## RESULT 20

US-10-092-302-951  
; Sequence 951, Application US/10092302  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT218C1  
; CURRENT APPLICATION NUMBER: US/10/092,302  
; CURRENT FILING DATE: 2002-03-07  
; NUMBER OF SEQ ID NOS: 1040  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 951  
; LENGTH: 262  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-302-951

## Query Match

Best Local Similarity 41.7%; Score 40; DB 7; Length 262;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FHCNFDVTRL 11  
| | | | | : |  
Db 51 FEECYFEDVTS 61

## RESULT 21

US-09-708-427-56000  
; Sequence 56000, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.

```

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56000
; LENGTH: 272
; TYPE: PRF
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..272
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..272
; OTHER INFORMATION: Ceres Seq. ID 1936468
US-09-708-427-56000
```

```

Query Match          41.7%; Score 40; DB 6; Length 272;
Best Local Similarity 46.2%; Pred. No. 43;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
Oy 1 FEHCNFDVTTTL 13
    1:||||: :1
Db 37 FKHCNFSFVROL 49
```

```

RESULT 22
US-09-708-427-55999
; Sequence 55999, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55999
; LENGTH: 299
; TYPE: PRF
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..299
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..299
; OTHER INFORMATION: Ceres Seq. ID 1936467
US-09-708-427-55999
```

```

Query Match          41.7%; Score 40; DB 6; Length 299;
Best Local Similarity 46.2%; Pred. No. 47;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
Oy 1 FEHCNFDVTTTL 13
    1:||||: :1
Db 64 FKHCNFSFVROL 76
```

```

RESULT 23
PCT-US02-03987-15325
; Sequence 15325, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for identifying the Target of a Compound which Inhibits
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITTRA.028VPC
```

```

; CURRENT APPLICATION NUMBER: PCT/US02/03987
; CURRENT FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15325
; LENGTH: 394
; TYPE: PRF
; ORGANISM: Candida albicans
PCT-US02-03987-15325
```

```

Query Match          41.7%; Score 40; DB 1; Length 394;
Best Local Similarity 53.8%; Pred. No. 65;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
Oy 5 NFNDVTTTLENE 17
    |||:|: |||
Db 349 NFNNVSAELKERE 361
```

```

RESULT 24
US-10-032-585-7381
; Sequence 7381, Application US/10032585
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7381
; LENGTH: 394
; TYPE: PRF
; ORGANISM: Candida albicans
US-10-032-585-7381
```

```

Query Match          41.7%; Score 40; DB 7; Length 394;
Best Local Similarity 53.8%; Pred. No. 65;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
Oy 5 NFNDVTTTLENE 17
    |||:|: |||
Db 349 NFNNVSAELKERE 361
```

```

RESULT 25
US-10-072-851-15325
; Sequence 15325, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for identifying the Target of a Compound which Inhibit
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITTRA.028A
```

```

; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 15325
; LENGTH: 394
; TYPE: PRN
; ORGANISM: Candida albicans
US-10-072-851-15325

```

```

Query Match
Best Local Similarity 41.7%; Score 40; DB 7; Length 394;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
OY 5 NNDVTTRLRENE 17
Db 349 NNNVSALERE 361

```

Search completed: April 2, 2002, 09:23:05  
 Job time: 194 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:19:51 ; Search time 12.43 Seconds  
(without alignments)  
30.777 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58  
Perfect score: 96  
Sequence: 1 FEHCNFNDVTRLRENE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	103	1	US-08-271-562-1
2	96	100.0	103	1	US-08-087-007-3
3	96	100.0	103	2	US-08-696-777-1
4	96	100.0	103	3	US-08-483-433-3
5	96	100.0	103	5	PCT-US92-05920-3
6	96	100.0	105	4	US-09-591-435-12
7	96	100.0	128	6	5179198-1
8	96	100.0	128	6	5521296-1
9	81	84.4	121	4	US-09-591-435-13
10	81	84.4	121	4	US-08-663-566-15
11	41	42.7	1162	2	US-08-023-610-15
12	41	42.7	1162	2	US-08-288-065-15
13	41	42.7	1162	2	US-08-362-240-15
14	41	42.7	1162	5	PCT-US95-10245-15
15	41	42.7	1162	5	PCT-US95-10245-15
16	40	41.7	217	1	US-08-277-231A-2
17	40	41.7	217	1	US-08-277-231A-12
18	40	41.7	217	2	US-08-473-750-1
19	40	41.7	217	2	US-08-473-750-5
20	40	41.7	217	2	US-08-477-326-1
21	39.5	41.1	217	2	US-08-477-326-5
22	38	39.6	278	4	US-09-257-490-15
23	37	38.5	278	4	US-08-392-828C-4
24	37	38.5	278	3	US-09-330-945-4
25	36.5	38.5	586	2	US-08-630-822A-70
26	36.5	38.0	486	2	US-09-005-069-70
27	36.5	38.0	491	3	US-08-889-841B-8
					Sequence 10, Appl

28	36	37.5	206	4	US-09-156-580-2	Sequence 2, Appl1
29	36	37.5	298	1	US-08-393-985-10	Sequence 10, Appl
30	36	37.5	300	4	US-09-433-598-2	Sequence 2, Appl1
31	36	37.5	312	4	US-08-821-872-2	Sequence 2, Appl1
32	36	37.5	364	1	US-08-680-726A-56	Sequence 56, Appl
33	36	37.5	364	4	US-09-092-409-56	Sequence 56, Appl
34	36	37.5	421	2	US-08-576-626A-53	Sequence 53, Appl
35	36	37.5	469	3	US-08-985-335-5	Sequence 5, Appl1
36	36	37.5	469	3	US-08-985-335-9	Sequence 5, Appl1
37	36	37.5	469	4	US-09-410-372-5	Sequence 9, Appl1
38	36	37.5	469	4	US-09-410-372-9	Sequence 9, Appl1
39	36	37.5	629	4	US-09-300-909-19	Sequence 31, Appl
40	36	37.5	1114	2	US-08-576-626A-31	Sequence 31, Appl
41	36	37.5	1233	4	US-09-194-613-5	Sequence 2, Appl1
42	36	37.5	1694	1	US-08-494-168-2	Sequence 2, Appl1
43	35.5	37.0	455	3	US-08-889-841B-46	Sequence 46, Appl
44	35.5	37.0	474	3	US-08-889-841B-36	Sequence 36, Appl
45	35.5	37.0	474	3	US-08-889-841B-39	Sequence 39, Appl
46	35.5	37.0	536	2	US-08-392-806A-2	Sequence 2, Appl1
47	35.5	37.0	536	4	US-09-257-490-2	Sequence 2, Appl1
48	35.5	37.0	600	2	US-08-392-806A-4	Sequence 4, Appl1
49	35.5	37.0	600	4	US-09-257-490-4	Sequence 4, Appl1
50	35.5	37.0	605	2	US-08-472-666-1	Sequence 1, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-271-562-1  
; Sequence 1, Application US/08271562  
; Patent No. 5573940  
; GENERAL INFORMATION:  
; APPLICANT: Sims, Peter J.  
; APPLICANT: Bothwell, Alfred L.M.  
; TITLE OF INVENTION: Genetic Inhibition of Complement  
; TITLE OF INVENTION: Mediated Inflammatory Response  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Kilpatrick & Cody  
; STREET: 1100 Peachtree Street, Suite 2800  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: U.S.  
; ZIP: 30309-4530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/271,562  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/729926  
; FILING DATE: 15-JUL-1991  
; APPLICATION NUMBER: US 07/365199  
; FILING DATE: 04-AUG-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: OMRF 112CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404) 815-6508  
; TELEFAX: (404) 815-6555  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 103 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

;;  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE: internal  
; ORGANISM: Homo sapiens  
; TISSUE TYPE: Blood  
; CELL TYPE: Erythrocyte  
US-08-271-562-1

Query Match  
Best Local Similarity 100.0%; Score 96; DB 1; Length 103;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFNDVTTRLRENE 17  
DB 42 FEHCNFNDVTTRLRENE 58

## RESULT 2

US-08-087-007-3  
; Sequence 3, Application US/08087007  
; Patent No. 5705732  
; Patent No. 5705732 5684223

## GENERAL INFORMATION:

APPLICANT: Sims, Peter J.  
APPLICANT: Bothwell, Alfred L.M.  
APPLICANT: Elliott, Eileen A.  
APPLICANT: Flavell, Richard A.  
APPLICANT: Madril, Joseph  
APPLICANT: Rollins, Scott  
APPLICANT: Bell, Leonard  
TITLE OF INVENTION: Squinto, Stephen  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 1100 Peachtree Street, Suite 2800  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: U.S.

ZIP: 30309-4530

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 19930701  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMR135  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-815-6500  
TELEFAX: 404-815-6555  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 103 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
CLONE: CD59  
US-08-087-007-3

Query Match  
Best Local Similarity 100.0%; Score 96; DB 1; Length 103;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFNDVTTRLRENE 17  
DB 42 FEHCNFNDVTTRLRENE 58

## RESULT 3

US-08-696-777-1  
; Sequence 1, Application US/08696777  
; Patent No. 5955441

## GENERAL INFORMATION:

APPLICANT: Sims, Peter J.  
APPLICANT: Bothwell, Alfred L.M.  
TITLE OF INVENTION: Genetic Inhibition of Complement Mediated  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2800 One Atlantic Center, 1201 West Peachtree Street  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA

ZIP: 30309-3450

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 13-AUG-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMR 112cipd1v  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404)-873-8794  
TELEFAX: (404)-873-8795  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 103 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-696-777-1

Query Match  
Best Local Similarity 100.0%; Score 96; DB 2; Length 103;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFNDVTTRLRENE 17  
DB 42 FEHCNFNDVTTRLRENE 58

## RESULT 4

US-08-483-433-3  
; Sequence 3, Application US/08483433  
; Patent No. 6100443

## GENERAL INFORMATION:

APPLICANT: Sims et al.  
TITLE OF INVENTION: Universal Donor Cells  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst

```
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,433
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/087,007
FILING DATE: July 1, 1993
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/906,394
FILING DATE: June 29, 1992
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF135c1p2 div
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-873-8794
TELEFAX: 404-873-9794
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: CD59
US-08-483-433-3

Query Match 100.0%; Score 96; DB 3; Length 103;
Best Local Similarity 100.0%; Pred. No. 4,8e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FEHCNFDVTRRLRENE 17
Db 42 FEHCNFDVTRRLRENE 58
```

```
RESULT 5
PCT-US92-05920-3
Sequence 3, Application PC/TUS9205920
GENERAL INFORMATION:
APPLICANT: Sims, Peter J.
APPLICANT: Bothwell, Alfred L.M.
APPLICANT: Elliott, Eileen A.
APPLICANT: Flavell, Richard A.
APPLICANT: Maertl, Joseph
APPLICANT: Rollins, Scott
APPLICANT: Squinto, Stephen
TITLE OF INVENTION: Universal Donor Cells
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street, Suite 2800
```

```
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/05920
FILING DATE: 19920714
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF135
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6500
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: CD59
PCT-US92-05920-3
```

```
Query Match 100.0%; Score 96; DB 5; Length 103;
Best Local Similarity 100.0%; Pred. No. 4,8e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FEHCNFDVTRRLRENE 17
Db 42 FEHCNFDVTRRLRENE 58
```

```
RESULT 6
US-09-591-435-12
Sequence 12, Application US/09591435
Patent No. 6280953
GENERAL INFORMATION:
APPLICANT: MESSIER, WALTER
APPLICANT: SIKELA, JAMES M
TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AND POLYPEPTIDE
SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL
CONDITIONS
FILE REFERENCE: GENO.200.2
CURRENT APPLICATION NUMBER: US/09/591,435
CURRENT FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 09/591,435
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 09/240,915
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: 60/073,263
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/098,987
PRIOR FILING DATE: 1998-09-02
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 12
LENGTH: 105
TYPE: PRT
ORGANISM: Homo sapiens
```

US-09-591-435-12

Query Match  
Best Local Similarity 100.0%; Score 96; DB 4; Length 105;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNNDVTTRLRENE 17  
Db 43 FEHCNNDVTTRLRENE 59

RESULT 7  
5179198-1  
; Patent No. 5179198  
; APPLICANT: OKADA, HIDECHIKA; OKADA, NORIKO; NAGAMI, YOICHI;  
; TAKASHI, KAZUHIRO; TAKIZAWA, HISAO; KONDO, JUN  
; TITLE OF INVENTION: GLYCOPROTEIN AND GENE CODING THEREFOR  
; NUMBER OF SEQUENCES: 17  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/376,828  
; FILING DATE: 07-JUL-1989  
; SEQ ID NO: 1:  
; LENGTH: 128  
5179198-1

Query Match  
Best Local Similarity 100.0%; Score 96; DB 6; Length 128;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNNDVTTRLRENE 17  
Db 67 FEHCNNDVTTRLRENE 83

RESULT 8  
5521296-1  
; Patent No. 5521296  
; APPLICANT: OKADA, HIDECHIKA; OKADA, NORIKO; NAGAMI, YOICHI;  
; TAKAHASHI, KAZUHIRO; TAKIZAWA, HISAO; KONDO, JUN  
; TITLE OF INVENTION: GLYCOPROTEIN AND GENE CODING THEREFOR  
; NUMBER OF SEQUENCES: 12  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/739,211  
; FILING DATE: 01-AUG-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 376,828  
; FILING DATE: 07-JUL-1989  
; SEQ ID NO: 1:  
; LENGTH: 128  
5521296-1

Query Match  
Best Local Similarity 100.0%; Score 96; DB 6; Length 128;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNNDVTTRLRENE 17  
Db 67 FEHCNNDVTTRLRENE 83

RESULT 9  
US-09-591-435-13  
; Sequence 13; Application US/09591435  
; Patent No. 6280953  
; GENERAL INFORMATION:  
; APPLICANT: MESSIER, WALTER  
; APPLICANT: SIKELA, JAMES M  
; TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AND POLYPEPTIDE  
; TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL  
; TITLE OF INVENTION: AND MEDICAL CONDITIONS

FILE REFERENCE: GENO.200.2  
; CURRENT APPLICATION NUMBER: US/09/591,435  
; CURRENT FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 09/591,435  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 09/240,915  
; PRIOR FILING DATE: 1999-01-29  
; PRIOR APPLICATION NUMBER: 60/073,263  
; PRIOR FILING DATE: 1998-01-30  
; PRIOR APPLICATION NUMBER: 60/098,987  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 121  
; TYPE: PRP  
; ORGANISM: Pan troglodytes  
US-09-591-435-13

Query Match  
Best Local Similarity 84.4%; Score 81; DB 4; Length 121;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EHCNFKDITRLRENE 17  
Db 44 EHCNFKDITRLRENE 59

RESULT 10  
US-08-663-566A-15  
; Sequence 15; Application US/08663566A  
; Patent No. 5853733  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D  
; APPLICANT: Macdonald, Richard D  
; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: John P. White  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/663,566A  
; FILING DATE: June 13, 1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P  
; REGISTRATION NUMBER: 28,678  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)278-0400  
; TELEFAX: (212)391-0526  
; TELETYPE: 422523  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1162 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-663-566A-15

Query Match  
Best Local Similarity 42.7%; Score 41; DB 2; Length 1162;  
Matches 75.0%; Pred. No. 97;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 HCNFNDVT 10  
|||||  
Db 101 HCNFSDTT 108

RESULT 11  
US-08-023-610-15  
; Sequence 15, Application US/08023610  
; Patent No. 5928648  
; GENERAL INFORMATION:  
; APPLICANT: Cochran Ph.D., Mark D  
; APPLICANT: Macdonald Ph.D., Richard D  
; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys  
; TITLE OF INVENTION: and Uses Thereof  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John P. White  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/023,610  
; FILING DATE: February 26, 1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White Esq, John P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)977-9550  
; TELEFAX: (212)664-0525  
; TELEX: 422523  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1162 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-023-610-15

Query Match 42.7%; Score 41; DB 2; Length 1162;  
Best Local Similarity 75.0%; Pred. No. 97;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 HCNFNDVT 10  
|||||  
Db 101 HCNFSDTT 108

RESULT 12  
US-08-288-065A-15  
; Sequence 15, Application US/0828065A  
; Patent No. 5961962  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D  
; APPLICANT: Macdonald, Richard D  
; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys S-  
; TITLE OF INVENTION: HVT-050 and Uses Thereof  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John P. White  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA

ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/288,065A  
; FILING DATE: Aug-09-94  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)278-0400  
; TELEFAX: (212)391-0526  
; TELEX: 422523  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1162 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-288-065A-15

Query Match 42.7%; Score 41; DB 2; Length 1162;  
Best Local Similarity 75.0%; Pred. No. 97;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 HCNFNDVT 10  
|||||  
Db 101 HCNFSDTT 108

RESULT 13  
US-08-362-240A-15  
; Sequence 15, Application US/08362240A  
; Patent No. 5965138  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D  
; APPLICANT: Junker, David  
; APPLICANT: Wild, Martha A  
; TITLE OF INVENTION: Recombinant Herpesvirus and Uses Thereof  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John P. White  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/362,240A  
; FILING DATE: Dec-22-94  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P  
; REGISTRATION NUMBER: 28,678  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)278-0400  
; TELEFAX: (212)391-0526  
; TELEX: 422523  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1162 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-362-240A-15

Query Match  
Best Local Similarity 42.7%; Score 41; DB 2; Length 1162;  
Matches 6; Conservative 1; Pred. No. 97;

QY 3 HCNFNDVT 10  
Db 101 HCNFSDT 108

## RESULT 14

PCT-US95-10245-15  
Sequence 15, Application PC/TUS9510245  
GENERAL INFORMATION:  
APPLICANT: SYNTRO CORPORATION  
TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys And Uses Thereof  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/10245  
FILING DATE: 09-AUG-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
REGISTRATION NUMBER: 28,678  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)278-0400  
TELEFAX: (212)391-0526  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1162 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-10245-15

Query Match  
Best Local Similarity 42.7%; Score 41; DB 5; Length 1162;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 3 HCNFNDVT 10  
Db 101 HCNFSDT 108

## RESULT 15

US-08-277-231A-2  
Sequence 2, Application US/08277231A  
Patent No. 5643725  
GENERAL INFORMATION:  
APPLICANT: Green, Bruce A.  
TITLE OF INVENTION: Sequence and Analysis of LKP Pili  
Patent No. 5643725  
TITLE OF INVENTION: Structural Genes and The LKP Pili Operon of No. 5643725typable  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02173

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/277,231A  
FILING DATE: 19-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Alice O.  
REGISTRATION NUMBER: 32,542  
REFERENCE/DOCKET NUMBER: ACC94-02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 217 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-277-231A-2

Query Match  
Best Local Similarity 41.7%; Score 40; DB 1; Length 217;  
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;  
QY 1 FEHCNNDVTTRLENE 17  
Db 156 FTHSNDVATQOTVVK 172

## RESULT 16

US-08-277-231A-12  
Sequence 12, Application US/08277231A  
Patent No. 5643725  
GENERAL INFORMATION:  
APPLICANT: Green, Bruce A.  
TITLE OF INVENTION: Sequence and Analysis of LKP Pili  
Patent No. 5643725  
TITLE OF INVENTION: Structural Genes and The LKP Pili Operon of No. 5643725typable  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/277,231A  
FILING DATE: 19-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Alice O.  
REGISTRATION NUMBER: 32,542  
REFERENCE/DOCKET NUMBER: ACC94-02  
TELECOMMUNICATION INFORMATION:



```

; TITLE OF INVENTION: Structural Gene and the LKP Pili Operon of No. 5968769typable
; NUMBER OF INVENTION: Haemophilus Influenzae
; CORRESPONDENCE ADDRESS: 21
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; APPLICATION NUMBER: US/08/477,326
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/277,231
; FILING DATE: July 19, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Alice O.
; REGISTRATION NUMBER: 33,542
; REFERENCE/DOCKET NUMBER: ACC94-02A
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-477-326-1

Query Match
Best Local Similarity 41.7%; Score 40; DB 2; Length 217;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 FEHCFNDVTRLRNE 17
Db 156 FTSHNDVATQOTVKN 172

RESULT 20
; US-08-477-326-5
; Sequence 5, Application US/08477326
; Patent No. 5968769
; GENERAL INFORMATION:
; APPLICANT: Green, Bruce A.
; APPLICANT: Brinton, Jr., Charles C.
; TITLE OF INVENTION: Sequence and Analysis of LKP Pili
; Patent No. 5968769
; TITLE OF INVENTION: Structural Gene and the LKP Pili Operon of No. 5968769typable
; NUMBER OF INVENTION: Haemophilus Influenzae
; CORRESPONDENCE ADDRESS: 21
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/477,326
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/277,231
; FILING DATE: July 19, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Alice O.
; REGISTRATION NUMBER: 33,542
; REFERENCE/DOCKET NUMBER: ACC94-02A
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-477-326-5

Query Match
Best Local Similarity 41.7%; Score 40; DB 2; Length 217;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 FEHCFNDVTRLRNE 17
Db 156 FTSHNDVATQOTVKN 172

RESULT 21
; US-09-257-490-15
; Sequence 15, Application US/09257490A
; Patent No. 6248328
; GENERAL INFORMATION:
; APPLICANT: Dietrich, Ursula
; APPLICANT: Von Briesen, Hagen
; APPLICANT: Grez, Manuel
; APPLICANT: Rubsamen-Waigmann, Helga
; TITLE OF INVENTION: HIV-1 virus isolates of a subtype and its differential
; TITLE OF INVENTION: diagnostics, a vaccine against HIV-1 virus infections
; TITLE OF INVENTION: of this subtype and method of producing same, use of
; FILE REFERENCE: 10496/P58512051
; CURRENT APPLICATION NUMBER: US/09/257,490A
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 15
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Simian immunodeficiency virus
; FEATURE:
; OTHER INFORMATION: SIVcpz
; US-09-257-490-15

Query Match
Best Local Similarity 41.1%; Score 39.5; DB 4; Length 602;
Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

QY 2 EHCNFDVTRLRNE 17
Db 45 KNCSEF-VTTELKDK 59

RESULT 22
; US-08-392-828C-4
; Sequence 4, Application US/08392828C
; Patent No. 5795962
; GENERAL INFORMATION:
; APPLICANT: IWANAGA, SADAOKI

```



APPLICANT: MOTA, TATSUSHI  
APPLICANT: SEKI, NORIAKI  
APPLICANT: ODA, TOSHIO  
TITLE OF INVENTION: NOVEL POLYPEPTIDE AND DNA ENCODING  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
ADDRESSEE: THIBEAULT  
STREET: 53 STATE STREET  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/392,828C  
FILING DATE: 28-FEB-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: CAMPBELL, PAULA A  
REGISTRATION NUMBER: 32,503  
REFERENCE/DOCKET NUMBER: FJN-033  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7100  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 278 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-392-828C-4

Query Match 39.6%; Score 38; DB 1; Length 278;  
Best Local Similarity 50.0%; Pred. No. 63;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 EHCNFNDVTTRL 13  
Db 5 KHCGRPVITRI 16

RESULT 23  
US-09-330-945-4  
Sequence 4, Application US/09330945  
Patent No. 6077946  
GENERAL INFORMATION:  
APPLICANT: IWANAGA, SADRAKI  
APPLICANT: MOTA, TATSUSHI  
APPLICANT: SEKI, NORIAKI  
APPLICANT: ODA, TOSHIO  
TITLE OF INVENTION: DNA ENCODING HORSESHOE CRAB  
TITLE OF INVENTION: AMEBOCYTE LYSATE FACTOR G SUBUNIT A  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
ADDRESSEE: THIBEAULT, LLP  
STREET: 125 HIGH STREET  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/330,945  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION NUMBER: 09/119,995  
APPLICATION NUMBER: 09/119,995  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: PITCHER, EDMUND R  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: FJN-032DV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7100  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 278 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-330-945-4

Query Match 39.6%; Score 38; DB 3; Length 278;  
Best Local Similarity 50.0%; Pred. No. 63;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 EHCNFNDVTTRL 13  
Db 5 KHCGRPVITRI 16

RESULT 24  
US-08-630-822A-70  
Sequence 70, Application US/08630822A  
Patent No. 5840695  
GENERAL INFORMATION:  
APPLICANT: FRANK, GLENN R.  
APPLICANT: HUNTER, SHIRLEY WU  
APPLICANT: WALLENFELS, LYNDIA  
TITLE OF INVENTION: NOVEL ECTOPARASITE SALLIYA PROTEINS  
TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross P.C.  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.A.  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/630,822A  
FILING DATE: 11-APR-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CONNELL, GARY J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-17-C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 586 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:

Tue Apr 2 09:51:55 2002

us-09-020-393b-3\_copy\_42\_58.open.ra1

NAME/KEY: Xaa = any amino acid  
LOCATION: 379  
US-08-630-822A-70

Search completed: April 2, 2002, 09:20:32  
Job time: 41 sec

Query Match  
Best Local Similarity 38.5%; Score 37; DB 2; Length 586;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
OY 2 EHCNENDVT 10  
Db 578 EHCQYPDIT 586

RESULT 25  
US-09-005-069-70  
; Sequence 70, Application US/09005069  
; Patent No. 5932470  
; GENERAL INFORMATION:  
; APPLICANT: FRANK, GLENN R.  
; APPLICANT: HUNTER, SHIRLEY WU  
; APPLICANT: WALLBENELS, LYNDIA  
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS  
; NUMBER OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross P.C.  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: U.S.A.  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/005,069  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/630,822  
; FILING DATE: 11-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CONNELL, GARY J  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-17-C3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 863-9700  
; TELEFAX: (303) 863-0223  
; INFORMATION FOR SEQ ID NO: 70:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 586 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Xaa = any amino acid  
; LOCATION: 379  
US-09-005-069-70

Query Match  
Best Local Similarity 38.5%; Score 37; DB 2; Length 586;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
OY 2 EHCNENDVT 10  
Db 578 EHCQYPDIT 586



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:01:52 ; Search time 37.51 Seconds  
(without alignments)  
54.831 Million cell updates/sec

Title: US-09-020-393b-18

Perfect score: 145

Sequence: 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR:68:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	142	97.9	559	1 C9HU	complement C9 prec
2	52	35.9	3237	2 T37964	probable cytoplasm
3	49	33.8	135	2 PN0494	transkeolase (EC
4	49	33.8	524	2 A29677	probable membrane
5	46.5	32.1	1002	2 T19226	hypothetical prote
6	46	31.7	254	2 T30493	hypothetical prote
7	46	31.7	254	2 E84826	hypothetical prote
8	46	31.7	292	2 C70421	conserved hypotet
9	46	31.7	357	2 S27909	hypothetical prote
10	46	31.7	357	2 J02174	hypothetical prote
11	46	31.7	930	2 JX0368	inter-alpha-trypsi
12	46	31.7	1372	2 T25933	hypothetical prote
13	46	31.7	2144	2 T21712	hypothetical prote
14	45.5	31.4	661	2 C81822	topoisomerase IV s
15	45.5	31.4	661	2 D81055	topoisomerase IV c
16	45	31.0	358	2 T44333	hypothetical prote
17	45	31.0	440	2 C84265	adenylosuccinate s
18	45	31.0	732	2 T18567	hypothetical prote
19	45	31.0	895	2 T49010	hypothetical prote
20	45	31.0	1189	2 T30319	liian-Aal retrorran
21	44.5	30.7	884	2 T18649	hypothetical prote
22	44	30.3	241	2 T26909	hypothetical prote
23	44	30.3	261	2 A86763	cobyrinic acid synth
24	44	30.3	535	2 F81847	dihydrolipeamide S
25	44	30.3	535	2 F81094	pyruvate dehydroge
26	44	30.3	560	2 S42158	KREI1 protein - ye
27	44	30.3	659	2 A81862	transkeolase (EC
28	44	30.3	659	2 B81082	transkeolase NM1
29	44	30.3	668	2 A46013	coagulation factor

30	44	30.3	678	2 A71287	probable cytoplasm
31	44	30.3	679	2 S54299	transkeolase (EC
32	44	30.3	944	1 S48821	probable membrane
33	44	30.3	971	2 E96794	hypothetical prote
34	43.5	30.0	117	2 H75165	hypothetical prote
35	43.5	30.0	247	2 T52511	related to lysopho
36	43.5	30.0	341	2 S64618	hypothetical prote
37	43.5	30.0	739	2 F86337	hypothetical prote
38	43.5	30.0	884	2 E86244	unknown protein, 4
39	43.5	30.0	2550	2 B53435	vesicular transpor
40	43	29.7	201	2 A47096	response regulator
41	43	29.7	257	2 C64442	hypothetical prote
42	43	29.7	305	2 F64481	hypothetical prote
43	43	29.7	309	2 T09564	glutaminyl-peptide
44	43	29.7	311	2 S52498	GPM2 protein - yea
45	43	29.7	323	2 T26536	hypothetical prote

#### ALIGNMENTS

RESULT 1  
C9HU  
complement C9 precursor [validated] - human  
C:Species: Homo sapiens (man)  
C:Date: 27-Nov-1985 #sequence.revision 17-Nov-2000 #text.change 17-Nov-2000  
C:Accession: A59363; I52400; A91029; A94019; S68647; A59364; A03208  
R:Maraziti, D.; Eggertsen, G.; Fey, G.H.; Stanley, K.K.  
unpublished results, 1988, cited by GenBank  
A:Description: Relationships between the gene and protein structure in human compleme  
A:Reference number: A59363  
A:Accession: A59363  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-559 <MAR1>  
A:Cross-references: GB:X02176; NID:929580; PIDN:CAA26117.1; PID:929581  
R:Maraziti, D.; Eggertsen, G.; Fey, G.H.; Stanley, K.K.  
Biochemistry 27, 6529-6534, 1988  
A:Title: Relationships between the gene and protein structure in human complement com  
A:Reference number: I52400; MUID:89118250  
A:Accession: I52400  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 62-159 <MAR2>  
A:Cross-references: GB:X02833; NID:9179727; PIDN:AAA51890.1; PID:9179728  
R:Stanley, K.K.; Kocher, H.P.; Luzzio, J.P.; Jackson, P.; Tschopp, J.  
EMBO J. 4, 375-382, 1985  
A:Title: The sequence and topology of human complement component C9.  
A:Reference number: A91029; MUID:85257464  
A:Accession: A91029  
A:Molecule type: mRNA  
A:Residues: 1-559 <STA>  
A:Cross-references: GB:X02176; NID:929580  
R:Disclipio, R.G.; Gehring, M.R.; Podack, E.R.; Kan, C.C.; Hugli, T.E.; Fey, G.H.  
Proc. Natl. Acad. Sci. U.S.A. 81, 7299-7302, 1984  
A:Title: Nucleotide sequence of cDNA and derived amino acid sequence of human complem  
A:Reference number: A94019; MUID:85063778  
A:Accession: A94019  
A:Molecule type: mRNA  
A:Residues: 2-12, 'X', 14-16, 'X', 18-42, 'R', 44-313, 315-416, 'P', 418-559 <DIS>  
A:Cross-references: GB:K02766; NID:9179725; PIDN:AAA51889.1; PID:9179726  
R:Langweiler, S.; Schaller, J.; Rickli, B.E.  
FEBS Lett. 380, 8-12, 1996  
A:Title: Identification of disulfide bonds in the ninth component (C9) of human compl  
A:Reference number: S68647; MUID:96181657  
A:Accession: S68647  
A:Molecule type: protein  
A:Residues: 34-47, 52-59, 69-87, 'X', 89-93, 'T', 94-98, 106-113, 118-131, 136-145, 180-181, 'X'  
R:Witze, Schlomp, K.; Hobart, M.J.; Fernie, B.A.; Otten, A.; Wurzner, R.; Rittner, C.;  
Immunogenetics 48, 144-147, 1998  
A:Title: Heterogeneity in the genetic basis of human complement C9 deficiency.  
A:Reference number: A59364; MUID:98296010  
A:Note: submitted to Genbank, September 1996

A:Accession: A59364  
A:Status: preliminary: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 27-559 <MIT>  
A:Cross-references: GB:Y08545; NID:q1834472; PIDN:CAA69849.1; PID:q2258128  
R:Histsteenge, J.; Blommestein, M.; Hess, D.; Furmanek, A.; Mitroshnikov, O.  
J. Biol. Chem. 274, 32786-32794, 1999  
A:title: The four terminal components of the complement system are C-mannosylated on multiple sites  
A:reference number: A53562; MUID:20020247  
A:Note: identification and location of C-mannosylation sites by mass-spectroscopy and (1D)  
C:genetics:  
A:gene: GDB:c9  
A:cross-references: GDB:119738; OMTM:120940  
A:map position: 5p13-5p13  
C:complex: monomer in plasma; 10 to 16 chain multimer in transmembrane form  
C:function:  
A:description: in association with complement C5b-8 complex polymerizes to form a transmembrane pore  
A:pathway: complement pathway  
C:superfamily: complement C9; EGF homology; LDL receptor ligand-binding repeat homology; F1:21/Domains: complement pathway; cytolysis; glycoprotein; inflammation; membrane attack C  
F:22-265/Domains: signal sequence #status predicted <SIG>  
F:22-265/Domains: product: complement C9 #status predicted <MAT>  
F:41-95/Domains: thrombospondin type 1 repeat homology <C9A>  
F:101-134/Domains: LDL receptor ligand-binding repeat homology <R>  
F:266-559/Domains: complement C9b #status predicted <C9b>  
F:510-539/Domains: EGF homology <EGF>  
F:43-78, 54-57, 88-94, 142-181, 254-255, 380-405, 510-526, 513-528, 530-539/Dissulfide bonds: #st  
F:51/Modified site: 2' mannose-1-tryptophan (Trp) #status experimental  
F:265-266/Cleavage site: His-Gly (thrombin) #status predicted  
F:277-415/Binding site: carboxydrate (asn) (covalent) #status predicted

```

Query Match          97.98;  Score 142;  DB 1;  length 559;
Best Local Similarity 96.38;  Pred. No. 2,6e-14;
Matches 26;  Conservative 1;  Mismatches 0;  Indels 0;  Gaps 0;

QY      1  KCLGYHLDVSLAFSEISVGAERNKDDC 27
          :|||||:|||||:|||||:|||||:
Db      379  RCLGYHLDVSLAFSEISVGAERNKDDC 405

```

RESULT	2
probable ubiquitin ligase - fission yeast (schizosaccharomyces pombe)	
C:Species	Schizosaccharomyces pombe
C:Date:	03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession:	T37964
R:Devlin, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.	
submitted to the EMBL Data Library, September 1997	
A:Reference number:	22157
A:Accession:	T37964
A:Status:	Preliminary
A:Molecule type:	DNA
A:Residues:	1-3227 <DEV>
A:Cross-references:	EMBL:299531; PIDN:CA816714.1; GSPDB:GN00066; SPDB:SPAC19D5.04
A:Experimental source:	strain 97zh-; cosmid cl9D5
C:Genetics:	
A:Gene:	SPDB:SPAC19D5.04
A:Map position:	1

Query Match	35.98;	Score 52;	DB 2;	Length 3227;
Best Local Similarity	46.28;	Pred. NO. 25;		
Matches 12;	Conservative	4;	Mismatches	
			10;	Indels
				0;
QY	2	CLGHTDYSIAFEISVGAENKDDC	27	
		:   :   :   :   :   :   :		
Db	1339	CILYLEVLADSKRPDEFENSEDC	1364	

```

RESULT      3
PN0494
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - Cherry salmon (fragment)
N:Alternate names: poly ADP-ribose polymerase
C:Species: Oncorhynchus masou (cherry salmon)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #extl_change 17-Mar-1999
C:Accession: PN0494
R:Ozawa, Y.; Uchida, K.; Aml, Y.; Kushida, S.; Okada, N.; Miwa, M.
Biochem. Biophys. Res. Commun. 193, 119-125, 1993
A:Title: Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose) polymerase
A:Reference number: PN0494; MUID:93277538
A:Accession: PN0494
A:Molecule type: mRNA
A:Residues: 1-135 <OZA>
C:Comment: This zinc-finger protein plays a role in DNA repair, cell growth, and differentiation.
C:Superfamily: NAD+ ADP-ribosyltransferase
C:Keywords: DNA binding; glycosyltransferase; hexosyltransferase; NAD; nucleus; pentamer

```

RESULT 4  
A29677  
complement C9 precursor - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 16-Jul-1999  
C:Accession: A29677  
R:Stanley, K.K.; Herz, J.  
EMBO J. 6, 1951-1957, 1987  
A:Title: Topological mapping of complement component C9 by recombinant DNA techniques  
A:Reference number: A29677; MUID:88004404  
A:Accession: A29677  
A:Molecule type: mRNA  
A:Residues: 1-524 <STA>  
A:Cross-references: GB:X05475  
C:Superfamily: complement C9; EGF homology; LDL receptor ligand-binding repeat homolo  
E:19-73/Domain: complement alternate pathway; complement pathway; cytolysis; glycoprotein  
F:79-112/Domain: LDL receptor ligand-binding repeat homology <THR>  
F:489-518/Domain: EGF homology <EGF>

Query Match	33.8%	Score 49	DB 2	Length 524
Best Local Similarity	33.3%	Pred. No. 11		
Matches	10	Conservative	6	Mismatches
			10	Indels
				Gaps
QY	2	CLGHLDVSLA----	FSEISYCAEENKDDC	27
		: :	: : :	
Db	354	CLGFMDLRIPLDDDLKDA	SVTASVNAADGC	383

RESULT 5  
 T19226  
 hypothetical protein C13B4.1 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T19226  
 R:Smey, R.  
 submitted to the EMBL Data Library, November 1996  
 A:Reference number: Z19093  
 A:Accession: T19226  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1002 <MIL>  
 A:Cross-references: EMBL:Z81466; PIDN:CAB03875.2; GSPDB:GN00020; CESP:C13B4.1.







Query Match 31.4%; Score 45.5; DB 2; Length 661;  
 Best Local Similarity 34.5%; Pred. No. 49;  
 Matches 10; Conservative 5; Mismatches 9; Indels 5; Gaps 1;

QY 3 LGYHLD-----VSLAFSEISVGAEFNKDD 26  
 :| | : | | : | | : | | : | | :  
 Db 78 VGLHPEEGVPVVELVFTRLHAGKFNKDD 106

## RESULT 15

D81055  
 topoisomerase IV chain B NMB1682 [imported] - Neisseria meningitidis (strain MC58 serog  
 C;Species: Neisseria meningitidis  
 C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
 C;Accession: D81055  
 R;Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A  
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
 ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.  
 Science 287, 1809-1815, 2000  
 A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
 A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A;Reference number: A81000; MUID:20175755  
 A;Accession: D81055  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-661 <FET>  
 A;Cross-references: GB:AE002518; GB:AE002098; NID:g7226928; PIDN:AAF42030.1; PID:g722693  
 A;Experimental source: serogroup B, strain MC58  
 C;Genetics:  
 A;Gene: NMB1682  
 C;Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain B

Query Match 31.4%; Score 45.5; DB 2; Length 661;  
 Best Local Similarity 34.5%; Pred. No. 49;  
 Matches 10; Conservative 5; Mismatches 9; Indels 5; Gaps 1;

QY 3 LGYHLD-----VSLAFSEISVGAEFNKDD 26  
 :| | : | | : | | : | | : | | :  
 Db 78 VGLHPEEGVPVVELVFTRLHAGKFNKDD 106

Search completed: January 6, 2002, 09:03:01  
 Job time: 69 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:04:17 ; Search time 25.81 Seconds

(without alignments)  
38.355 Million cell updates/sec

Title: US-09-020-393b-18

Perfect score: 145  
Sequence: 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	142	97.9	559	1	CO9_HUMAN
2	61.5	42.4	547	1	CO9_HORSE
3	59	40.7	554	1	CO9_RAT
4	57	39.3	557	1	CO9_RABIT
5	49	33.8	135	1	PPOL_ONCMA
6	49	33.8	528	1	CO9_MOUSE
7	46	31.7	930	1	ITHD_HUMAN
8	44	30.3	165	1	MORE_STNP7
9	44	30.3	560	1	KR11_YEAST
10	44	30.3	668	1	F13B_MOUSE
11	44	30.3	677	1	CFPA_TREPA
12	44	30.3	679	1	TKTA_CRAPL
13	44	30.3	901	1	VGLE_GRCMV
14	44	30.3	944	1	YMH6_YEAST
15	43.5	30.0	341	1	YG5X_YEAST
16	43	29.7	305	1	YES5_METJA
17	43	29.7	311	1	PMG2_YEAST
18	43	29.7	330	1	YDC3_SCHPO
19	43	29.7	417	1	PGK_METJA
20	43	29.7	1070	1	PI1B_HUMAN
21	42.5	29.3	600	1	XKSL_YEAST
22	42	29.0	238	1	ATP6_SPRPN
23	42	29.0	310	1	YBR1_YEAST
24	42	29.0	329	1	YX9_BACHD
25	42	29.0	379	1	CYB_ASTPE
26	42	29.0	484	1	PURA_MAIZE
27	42	29.0	683	1	PIOL_SCHPO
28	42	29.0	758	1	YJ06_YEAST
29	42	29.0	1052	1	BU1B_MOUSE
30	42	29.0	1877	1	PCK5_MOUSE
31	41.5	28.6	281	1	Y246_MYCPN
32	41.5	28.6	435	1	NHG2_PSEPU
33	41.5	28.6	790	1	YDDB_ECOLI

RESULT	ID	CO9_HUMAN	STANDARD	PRT	559 AA.	ALIGNMENTS
34	41.5	28.6	986	1	EP1B_STREP	P30195 staphylococ
35	41	28.3	230	1	CRP_RAT	P48199 rattus norv
36	41	28.3	440	1	G3PT_MOUSE	O64467 mus musculu
37	41	28.3	481	1	6PGD_DROME	P41572 drosophila
38	41	28.3	481	1	6PGD_DROSI	P41573 drosophila
39	41	28.3	560	1	EMA5_HUMAN	O00341 homo sapien
40	41	28.3	667	1	TKT2_ECOLI	P33570 escherichia
41	41	28.3	901	1	CR14_MAIZE	O24585 zea mays (m
42	41	28.3	921	1	ITHD_PIG	P79263 sus scrofa
43	41	28.3	1070	1	PI1B_RAT	O92110 rattus norv
44	41	28.3	2647	1	ABP2_HUMAN	P21333 homo sapien
45	40.5	27.9	250	1	YOUT_CABEL	P34630 caenorhabdi

RESULT 1

CO9\_HUMAN

ID P02748;

AC P02748;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE COMPLEMENT COMPONENT C9 PRECURSOR.

DE C9.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=85257464; PubMed=4018030;

RA Stanley K.K., Kocher H.-P., Luzzo J.P., Jackson P., Tschopp J.;

RT "The sequence and topology of human complement component C9.";

RL EMBO J. 4:375-382(1985).

RN [2]

RP SEQUENCE OF 2-559 FROM N.A.

RX MEDLINE=85063778; PubMed=6095282;

RA Discipio R.G., Gehring M.R., Podack E.R., Kan C.C., Hugli T.E.,

RT "Nucleotide sequence of cDNA and derived amino acid sequence of human complement component C9.";

RL Proc. Natl. Acad. Sci. U.S.A. 81:7298-7302(1984).

RN [3]

RP SEQUENCE OF 62-159 FROM N.A.

RX MEDLINE=89118250; PubMed=3219351;

RA Maraziti D., Eggertsen G., Fey G.H., Stanley K.K.;

RT "Relationships between the gene and protein structure in human complement component C9.";

RL Biochemistry 27:6529-6534(1988).

RN [4]

RP 3D-STRUCTURE MODELLING OF MEMBRANE-SPANNING DOMAIN (MSB).

RX MEDLINE=90370039; PubMed=2395434;

RA Feitich M.C., Amiquet P., Guy R., Brunner J., Malzel J.V. Jr.,

RT Tschopp J.;

RT "Localization and molecular modelling of the membrane-inserted domain of the ninth component of human complement and perforin.";

RL Mol. Immunol. 27:589-602(1990).

RN [5]

RP CARBOHYDRATE-LINKAGE SITES.

RX MEDLINE=20020247; PubMed=10551839;

RA Hofsteenge J., Blommestein M., Hess D., Furmanek A., Miroshnichenko O.;

RT "The four terminal components of the complement system are C-mannosylated on multiple tryptophan residues.";

RL J. Biol. Chem. 274:32786-32794(1999).

CC -1- FUNCTION: C9 IS THE FINAL COMPONENT OF THE COMPLEMENT SYSTEM TO BE ADDED IN THE ASSEMBLY OF THE MEMBRANE ATTACK COMPLEX. IT IS ABLE TO ENTER LIPID BILAYERS, FORMING TRANSMEMBRANE CHANNELS.

CC -1- PTM: THROMBIN CLEAVES FACTOR C9 TO PRODUCE C9A AND C9B.

CC -1- SIMILARITY: TO COMPLEMENT FACTORS C6, C7, C8, AND TO PERFORIN.

CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.



FT DISULFID 530 539 BY SIMILARITY.  
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 451 451 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 547 AA; 62014 MW; 75E5FE961DE873B6 CRC64;

Query Match 42.4%; Score 61.5; DB 1; Length 547;  
 Best Local Similarity 44.4%; Pred. No. 0.062;  
 Matches 12; Conservative 8; Mismatches 6; Indels 1; Gaps 1;

OY 1 KCIGYHLDVSLAFSEISVGAEFNKDDC 27  
 DB 379 RCLGFNDVSLK-DKYEVTAKIDKNDK 404

RESULT 3  
 ID CO9\_RAT STANDARD; PRT; 554 AA.  
 AC Q62930;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE COMPLEMENT COMPONENT C9 PRECURSOR.  
 GN C9.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY;  
 RX MEDLINE=97355567; PubMed=9212048;  
 RA Lassiter H.A., Walz B.M., Wilson J.L., Jung E., Calisi C.R.,  
 RA Goldsmith L.J., Wilson R.A., Morgan B.P., Feldhoff R.C.,  
 RT "The administration of complement component C9 enhances the survival  
 of neonatal rats with Escherichia coli sepsis."  
 RL Pediatr. Res. 42:128-136(1997).  
 CC -1- FUNCTION: C9 IS THE FINAL COMPONENT OF THE COMPLEMENT SYSTEM TO  
 BE ADDED IN THE ASSEMBLY OF THE MEMBRANE ATTACK COMPLEX. IT IS  
 ABLE TO ENTER LIPID BILAYERS, FORMING TRANSMEMBRANE CHANNELS.  
 CC -1- SIMILARITY: TO COMPLEMENT FACTORS C6, C7, C8, AND TO PERFORIN.  
 CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: U49071; AAB38023.1; -  
 DR HSSP: P01130; ILDL.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR002172; LDL-recept.A.  
 DR InterPro: IPR001662; MAC-perforin.  
 DR InterPro: IPR000884; TSP1.  
 DR Pfam: PF00057; ldl\_recept\_a; 1.  
 DR Pfam: PF01823; MACPF; 1.  
 DR PRINTS: PR00764; COMPLEMENTC9.  
 DR SMART: SM00192; LDLa; 1.  
 DR SMART: SM00457; MACPF; 1.  
 DR SMART: SM00209; TSP1; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; FALSE\_NEG.  
 DR PROSITE: PS01209; ILDLA\_1; 1.  
 DR PROSITE: PS00068; ILDLA\_2; 1.  
 DR PROSITE: PS00279; MAC\_PERFORIN; 1.  
 DR PROSITE: PS00092; TSP1; 1.

KW Complement pathway; Complement alternate pathway; Glycoprotein;  
 KW Plasma; Membrane attack complex; Cytolysis; Transmembrane; Signal;  
 KW EGF-like domain.  
 FT SIGNAL 1 20  
 FT CHAIN 21 554  
 FT DOMAIN 40 96  
 FT DOMAIN 97 134  
 FT TRANSMEM 316 332  
 FT TRANSMEM 337 356  
 FT DOMAIN 515 549  
 FT SITE 267 268  
 FT DISULFID 99 110  
 FT DISULFID 105 123  
 FT DISULFID 117 132  
 FT DISULFID 382 413  
 FT DISULFID 519 535  
 FT DISULFID 522 537  
 FT DISULFID 539 548  
 FT CARBOHYD 261 261  
 FT CARBOHYD 409 409  
 FT CARBOHYD 423 423  
 SO SEQUENCE 554 AA; 62280 MW; 9C885F76A1275649 CRC64;

Query Match 40.7%; Score 59; DB 1; Length 554;  
 Best Local Similarity 42.4%; Pred. No. 0.16;  
 Matches 14; Conservative 4; Mismatches 9; Indels 6; Gaps 1;

OY 1 KCIGYHLDVSL-----AFSEISVGAEFNKDDC 27  
 DB 381 RCLGFNDVSLYTPLOTALGPSLTJANVHNSDC 413

RESULT 4  
 ID CO9\_RABIT STANDARD; PRT; 557 AA.  
 AC P48747;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE COMPLEMENT COMPONENT C9 PRECURSOR.  
 GN C9.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-34.  
 RC STRAIN=NEW ZEALAND WHITE; TISSUE=Liver;  
 RX MEDLINE=95181293; PubMed=753152;  
 RA Huesler T., Lockert D.H., Kaufman K.M., Sodez J.M., Sims P.J.,  
 RT "Chimers of human complement C9 reveal the site recognized by  
 complement regulatory protein CD59."  
 RL J. Biol. Chem. 270:3483-3486(1995).

CC -1- FUNCTION: C9 IS THE FINAL COMPONENT OF THE COMPLEMENT SYSTEM TO  
 BE ADDED IN THE ASSEMBLY OF THE MEMBRANE ATTACK COMPLEX. IT IS  
 ABLE TO ENTER LIPID BILAYERS, FORMING TRANSMEMBRANE CHANNELS.  
 CC -1- PTM: THROMBIN CLEAVES FACTOR C9 TO PRODUCE C9A AND C9B.  
 CC -1- SIMILARITY: TO COMPLEMENT FACTORS C6, C7, C8, AND TO PERFORIN.  
 CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: U20055; AAC48459.1; -  
 DR HSSP: P01130; IAUJ.

DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR002172; LDL\_recept\_A.  
 DR InterPro: IPR001862; MAC\_perforin.  
 DR Pfam: PF00084; TSP1.  
 DR Pfam: PF00057; ldl\_recept\_a; 1.  
 DR Pfam: PF01823; MACPF; 1.  
 DR Pfam: PF00090; tsp; 1; 1.  
 DR PRINTS: PRO0764; COMPLEMENTC9.  
 DR SMART: SM00192; LDLA; 1.  
 DR SMART: SM00457; MACPF; 1.  
 DR SMART: SM00209; TSP1; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS01209; LDLRA\_1; 1.  
 DR PROSITE: PS00068; LDLRA\_2; 1.  
 DR PROSITE: PS00279; MAC\_PERFORIN; 1.  
 DR PROSITE: PS00092; TSP1; 1.  
 KW Complement pathway; Complement alternate pathway; Glycoprotein;  
 KW EGF-like domain; Complement attack complex; Cytolysis; Transmembrane; Signal;  
 FT SIGNAL 1 21  
 FT CHAIN 1 21  
 FT DOMAIN 22 557 COMPLEMENT COMPONENT C9.  
 FT DOMAIN 99 98 TSP TYPE-1.  
 FT TRANSMEM 319 137 LDL-RECEPTOR CLASS A.  
 FT TRANSMEM 340 335 POTENTIAL.  
 FT DOMAIN 340 359 POTENTIAL.  
 FT DISULFID 316 550 EGF-LIKE.  
 FT DISULFID 101 113 BY SIMILARITY.  
 FT DISULFID 108 126 BY SIMILARITY.  
 FT DISULFID 120 135 BY SIMILARITY.  
 FT DISULFID 385 414 BY SIMILARITY.  
 FT DISULFID 520 536 BY SIMILARITY.  
 FT DISULFID 523 538 BY SIMILARITY.  
 FT DISULFID 540 549 BY SIMILARITY.  
 FT CARBOHD 261 261 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHD 282 282 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHD 424 424 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT SEQUENCE 557 AA; 62662 MW; FF4B5FEF8D1AB417 CRC64;

Query Match  
 Best Local Similarity 39.3%; Score 57; DB 1; Length 557;  
 Matches 12; Conservative 8; Mismatches 7; Indels 4; Gaps 1;

QY 1 KCLGYHLDSVLAFL---SEISVGAEPNKKDDC 27  
 DB 384 KCLGFPLDLSLNPGRKSGSLTNGQANKNKC 414

## RESULT 5

ID POL\_NCMA STANDARD; PRT: 135 AA.  
 AC 008824;

DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE POLY [ADP-RIBOSE] POLYMERASE (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+)-ADP-  
 GN ADPRT.  
 OS Oncorhynchus masou (Cherry salmon) (Masu salmon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Proteocephala; Salmoniformes; Salmonidae; Oncorhynchus.  
 NCBI\_TaxID=8020;

RP SEQUENCE FROM N.A.  
 RX MEDLINE=93277538; PubMed=8503897;  
 RA Osawa Y., Uchida K., Uchida M., Ami Y., Kushida S., Okada N.,  
 RT "Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)  
 RT polymerase from xenopus laevis and cherry salmon using heterologous  
 RT oligonucleotide consensus sequences";  
 RL Biochem. Biophys. Res. Commun. 193:119-125(1993).

CC CC  
 CC -1- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR  
 CC PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT  
 CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
 CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
 CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
 CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
 CC -1- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-RIBOSYL)[N-1]-ACCEPTOR =  
 CC NICOTINAMIDE + (ADP-D-RIBOSYL)[N-1]-ACCEPTOR.  
 CC -1- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN OVARY, OOCYTES,  
 CC AND BRAIN LOW IN LIVER.  
 CC -1- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
 CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
 CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
 CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
 CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
 CC -1- SIMILARITY: BELONGS TO THE PARP FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D13809; BAA02965.1; -  
 CC PIR: P0494; P0494.  
 CC HSSP: P26446; APAX.  
 CC DR P26446; APAX.  
 CC DR InterPro: IPR001290; PARP.  
 CC DR InterPro: IPR001510; ZnF-PARP.  
 CC Pfam: PF00644; PARP; 1.  
 CC DR PROSITE: PS00347; PARP\_ZN\_FINGER\_1; PARTIAL.  
 CC DR PROSITE: PS50064; PARP\_ZN\_FINGER\_2; PARTIAL.  
 CC KW Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
 FT ADP-ribosylation; Zinc-finger; Zinc.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 1  
 FT ACT\_SITE 135 135 NAD-BINDING.  
 FT NON\_TER 135 135 BY SIMILARITY.  
 SQ SEQUENCE 135 AA; 15411 MW; A60E8E9890EA2DC CRC64;

Query Match  
 Best Local Similarity 33.8%; Score 49; DB 1; Length 135;  
 Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 7 LDVSLAFSEISVGAEPNKKDD 26  
 DB 11 LDIEVAVSLKGAEDNKKD 30

## RESULT 6

ID CO9\_MOUSE STANDARD; PRT: 528 AA.  
 AC P06683;

DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE COMPLEMENT COMPONENT C9 (FRAGMENT).  
 GN C9.

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;

RP SEQUENCE FROM N.A.  
 RX MEDLINE=88004404; PubMed=2443347;  
 RA Stanley K.K., Herz J.;  
 RT "Topological mapping of complement component C9 by recombinant DNA  
 RT techniques suggests a novel mechanism for its insertion into target  
 RT membranes.";



```

DR EMBL: U43157; AAD05198.1; JOINED.
DR EMBL: U43158; AAD05198.1; JOINED.
DR EMBL: U43159; AAD05198.1; JOINED.
DR EMBL: U43160; AAD05198.1; JOINED.
DR EMBL: U43161; AAD05198.1; JOINED.
DR EMBL: U43162; AAD05198.1; JOINED.
DR MIM: 600564; -.
DR InterPro: IPR002035; WMFA.
DR Pfam: PF000092; VWA; 1.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS50234; WMFA; 1.
KW Serine protease inhibitor; Repeat; Signal; Multigene family;
KM Glycoprotein.
FT SIGNAL 1 28
FT CHAIN 29 661
FT PROPEP 662 668
FT CHAIN 689 930
FT DOMAIN 272 432
FT CARBOHYD 81 81
FT CARBOHYD 207 207
FT CARBOHYD 517 517
FT CARBOHYD 577 577
FT CARBOHYD 686 686
FT CARBOHYD 701 701
FT CARBOHYD 702 702
FT DISULFID 747 925
FT CONFLICT 85 85
FT CONFLICT 85 85
FT CONFLICT 114 114
FT CONFLICT 207 207
FT CONFLICT 221 221
FT CONFLICT 307 307
FT CONFLICT 322 322
FT CONFLICT 816 817
SQ SEQUENCE 930 AA; 103358 MW; 0E1929065FE4EB6A0 CRC64;

Query Match
Best Local Similarity 31.7%; Score 46; DB 1; Length 930;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 CUGYHLDVSLAFSE 15
DB 414 CLGGEFVSYAFLE 427

RESULT 8
MOE_SYNP7
ID MOE_SYNP7 STANDARD; PRT; 165 AA.
AC 056210;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE MOLYBDOTERIN [MPT] CONVERTING FACTOR, SUBUNIT 2 (MOLYBDENUM COFACTOR
DE BIOSYNTHESIS PROTEIN E) (MOLYBDOTERIN CONVERTING FACTOR LARGE
DE SUBUNIT).
GN MOAE.
OS Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=1140;
RN [1]
RP SEQUENCE FROM N.A.
RA Rubio L.M., Flores E., Herrero A.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CONVERSION OF MOLYBDOTERIN PRECURSOR 2 INTO
CC MOLYBDOTERIN REQUIRES TRANSFER OF TWO SULFUR ATOMS TO PRECURSOR 2
CC (TO GENERATE THE DITHIOLENE GROUP). THIS IS CATALYZED BY THE
CC CONVERTING FACTOR COMPOSED OF A SMALL AND LARGE SUBUNIT (BY
CC SIMILARITY).
CC -1- PATHWAY: MOLYBDENUM COFACTOR BIOSYNTHESIS.
CC -1- SUBUNIT: HETERODIMER OF MOAD AND MOAE.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X99625; CAA67947.1; -.
DR InterPro: IPR003448; MOEA.
DR Pfam: PF02391; MOEA; 1.
KW Molybdenum cofactor biosynthesis.
SQ SEQUENCE 165 AA; 18712 MW; AC016984E1D587FB CRC64;

Query Match
Best Local Similarity 30.3%; Score 44; DB 1; Length 165;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 2 CUGYHLDVSLAFSEISVGAEPKDD 26
DB 7 CDRHQIELSLAPRPLSAAEFCDD 31

RESULT 9
KR11_YEAST
ID KR11_YEAST STANDARD; PRT; 560 AA.
AC P32893;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE BETA-GLUCAN SYNTHESIS-ASSOCIATED PROTEIN KRE11 (KILLER TOXIN-
DE RESISTANCE PROTEIN 11).
GN KRE11 OR YGR166W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93216084; PubMed=8462845;
RA Brown J.L., Kosaczka Z., Jiang B., Bussey H.;
RT "A mutational analysis of killer toxin resistance in Saccharomyces
RT cerevisiae identifies new genes involved in cell wall (1-->6)-beta-
RT glucan synthesis."
RL Genetics 133:837-849(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=S288C;
RA MEDLINE=97435481; PubMed=9290212;
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT chromosome VII."
RL Yeast 13:1077-1090(1997).
RN [3]
RP FUNCTION: INVOLVED IN CELL WALL (1-->6)-BETA-GLUCAN SYNTHESIS.
CC KRE11 PROBABLY INTERACTS WITH KRB6. THEIR ROLE PRECEDES THAT OF
CC -1- PATHWAY: BIOSYNTHESIS OF BETA-GLUCANS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L10667; AA34727.1; -.
DR EMBL: Z72951; CAA97191.1; -.
DR PIR: S30784; S30784.
DR PIR: S42158; S42158.

```





```

DR EMBL:Z46647; CAAB86608.1; -.
DR HSP: P23254; IAT0.
DR InterPro: IPR000360; Transketolase.
DR Pfam: PF00456; transketolase_1.
DR PROSITE: PS00801; TRANSKETOLASE_1; 1.
DR PROSITE: PS00802; TRANSKETOLASE_2; 1.
KW transferase; Thiamine pyrophosphate; Multigene family.
SQ SEQUENCE 679 AA; 73130 MW; 1109092E136A345B CRC64;

Query Match
Best Local Similarity 30.3%; Score 44; DB 1; Length 679;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0

Qy 9 VSIAFSEISGAEKND 26
Db 139 VGLALAKKHGARINKP 156
||| : | : || : || : |

RESULT 13
VGLB_GPCMV STANDARD; PRT; 901 AA.
AC 069024;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE GLYCOPROTEIN B PRECURSOR.
GN GB OR UL55.
OS Guinea pig cytomegalovirus (strain 22122 / ATCC VR682) (GPCMV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.
NCBI_TaxID=103920;
[1]
SEQUENCE FROM N.A.
MEDLINE=94279138; PubMed=8009831;
Schleiss M.R.;
"Cloning and characterization of the guinea pig cytomegalovirus
glycoprotein B gene."
Virology 202;173-185(1994).
-- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: L25706; AAA43831.1; -.
DR InterPro: IPR000234; Glycoprot. B.
DR Pfam: PF00606; Glycoprotein_B.1.
DR Prodom: PD000693; Glycoprot_B.1.
KW glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 ?
FT CHAIN ? ?
FT DOMAIN ? 901 GLYCOPROTEIN B.
FT TRANSMEM 714 734 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 737 757 POTENTIAL.
FT DOMAIN 758 901 POTENTIAL.
FT CARBOHYD 53 53 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 60 60 N-LINKED (GLCNAC . . . ) (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC . . . ) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC . . . ) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC . . . ) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC . . . ) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC . . . ) (POTENTIAL).
FT CARBOHYD 372 372 N-LINKED (GLCNAC . . . ) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC . . . ) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC . . . ) (POTENTIAL).
FT CARBOHYD 436 406 N-LINKED (GLCNAC . . . ) (POTENTIAL).
FT CARBOHYD 457 436 N-LINKED (GLCNAC . . . ) (POTENTIAL).
FT CARBOHYD 537 537 N-LINKED (GLCNAC . . . ) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC . . . ) (POTENTIAL).

```

FT CARBOHYD 623 623 N-LINKED (GLCNAC.. ) (POTENTIAL).  
SQ SEQUENCE 901 AA; 102239 MW; 43ABCA/D45A197AA CRC64;

Query Match 30.3%; Score 44; DB 1; Length 901;  
Best Local Similarity 34.6%; Pred. No. 55;  
Matches 9; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

OY 1 KCLGYHLVSLAFSEISVGAFFNKDD 26  
Db 815 KDLSDVADTRVSSSSAGAGDFNEED 840

RESULT 14  
YMH6\_YEAST  
ID YMH6\_YEAST STANDARD; PRT; 944 AA.

AC Q03631;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE PUTATIVE 107.6 KDA TRANSCRIPTIONAL REGULATORY PROTEIN IN CPR3-HMG1  
DE INTERGENIC REGION.  
GN YML076C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID:4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C / AB972;  
RA Brown D., Bowman S., Barrell B.G., Rajandream M.A.;  
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: LOCATION: NUCLEAR (PROBABLE).  
CC -1- SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BIOMOLECULAR  
CC CLUSTER DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: Z46373; CAA86502.1; -  
DR SGD: S0004541; YML076C.  
DR InterPro: IPR001138; ZN2\_Cy6\_fungal.  
DR SMART: SM00066; GAL4; 1.  
DR PROSITE: PS00463; ZN2\_Cy6\_FUNGAL\_1; 1.  
DR PROSITE: PS50048; ZN2\_Cy6\_FUNGAL\_2; FALSE\_NEG.  
KW Hypothetical protein; Transcription regulation; DNA-binding; Zinc;  
KW Nuclear protein; Metal-binding.  
FT DNA\_BIND 76 109 ZN(2)-CYS(6), FUNGAL-TYPE.  
SQ SEQUENCE 944 AA; 107560 MW; 020A56745DF52CCC CRC64;

Query Match 30.3%; Score 44; DB 1; Length 944;  
Best Local Similarity 35.0%; Pred. No. 58;  
Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY 7 LDVSLAFSEISVGAFFNKDD 26  
Db 1 MDTOAITGVAVGKEINNND 20

RESULT 15  
YG5X\_YEAST  
ID YG5X\_YEAST STANDARD; PRT; 341 AA.  
AC P53336;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HYPOTHETICAL 38.5 KDA PROTEIN IN BGL2-ZUO1 INTERGENIC REGION.

GN YGR283C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID:4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C / FY1679;  
RX MEDLINE-97245295; Pubmed-9090054;

RA Volckaert G., Voet M., Robben J.,  
RT "Sequence analysis of a near-subtelomeric 35.4 kb DNA segment on the  
RT right arm of chromosome VII from Saccharomyces cerevisiae carrying  
RT the MAL1 locus reveals 15 complete open reading frames, including  
RT ZUO1, BGL2 and BIO2 genes and an ABC transporter gene."  
RL Yeast 13:251-259 (1997).  
CC -1- SIMILARITY: TO YEAST YMR310C.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: Z73068; CAA97314.1; -  
DR SGD: S0003515; YGR283C.  
DR InterPro: IPR003750; DUF171.  
DR Pfam: PF02598; DUF171; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 341 AA; 38546 MW; 3C91F8EF9C24EFC9 CRC64;

Query Match 30.0%; Score 43.5; DB 1; Length 341;  
Best Local Similarity 28.6%; Pred. No. 25;  
Matches 10; Conservative 8; Mismatches 8; Indels 9; Gaps 1;

OY 1 KCLGYHLVSLAFSEISVGAFFNKDD 26  
Db 216 KPLGYHMASTLNEVSEGYTKIYVWNSGDRHYDE 250

Search completed: January 6, 2002, 09:11:32  
Job time: 435 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:03:07 ; Search time 62.25 Seconds  
(without alignments)  
63.443 Million cell updates/sec

Title: US-09-020-393b-18  
Sequence: 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_17:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	40.7	567	11	Q62957
2	52	35.9	302	9	Q9E231
3	52	35.9	326	2	051948
4	52	35.9	326	2	0924K1
5	52	35.9	326	2	051944
6	52	35.9	327	2	0924L5
7	52	35.9	327	3	013834
8	47	32.4	241	5	0917L2
9	47	32.4	278	10	Q9SSX1
10	47	32.4	367	10	Q9SNG2
11	47	32.4	391	5	09V3Z2
12	46.5	32.1	1002	5	045247
13	46	31.7	203	12	Q9YMI5
14	46	31.7	254	10	Q9XEF0
15	46	31.7	292	2	067397
16	46	31.7	357	12	067692
17	46	31.7	376	2	087140
18	46	31.7	377	2	034236
19	46	31.7	644	4	Q9P190

20	46	31.7	1372	5	P91526
21	46	31.7	2144	5	062218
22	45.5	31.4	661	2	Q9JY97
23	45.5	31.4	661	2	Q9JY79
24	45	31.0	358	2	087162
25	45	31.0	440	1	Q9HQM6
26	45	31.0	732	5	Q18367
27	45	31.0	895	10	Q9LX29
28	45	31.0	1189	5	018558
29	45	31.0	2524	5	Q9GPA5
30	44.5	30.7	623	10	081082
31	44.5	30.7	909	5	Q17429
32	44	30.3	71	6	Q29226
33	44	30.3	160	6	Q9GLE7
34	44	30.3	241	5	062460
35	44	30.3	261	2	Q9CGJ1
36	44	30.3	311	10	Q9LJW1
37	44	30.3	380	2	Q9FCU8
38	44	30.3	423	8	Q9B896
39	44	30.3	424	8	Q9B8Z8
40	44	30.3	535	2	Q9JZ11
41	44	30.3	535	2	Q9JU07
42	44	30.3	571	2	Q9L4P0
43	44	30.3	659	2	Q9JTS0
44	44	30.3	659	2	Q9JTR1
45	44	30.3	662	2	Q99UD4

#### ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	567 AA.
Q62957	Q62957	01-NOV-1996 (TREMBLrel. 01, Created)		
AC	Q62957	01-NOV-1996 (TREMBLrel. 01, last sequence update)		
DT	Q62957	01-JUN-2001 (TREMBLrel. 17, last annotation update)		
DE	Q62957	COMPLEMENT COMPONENT C9 PRECURSOR.		
OS	Q62957	Rattus norvegicus (Rat).		
OC	Q62957	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Q62957	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.		
ON	Q62957	NCBI_TaxID=10116;		
OX	Q62957	[1]		
RP	Q62957	SEQUENCE FROM N.A.		
RC	Q62957	STRATIN-SPRAGUE DAWLEY.		
RA	Q62957	Hinchliffe S.J., den Berg C.W., Rushmere N.K., Morgan B.P.;		
RL	Q62957	Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.		
CC	Q62957	-1- SIMILARITY: TO LOW DENSITY LIPOPROTEIN (LDL) RECEPTOR CLASS A (LDLRA) DOMAIN.		
DR	Q62957	EMBL: U52948: AAA96528.1; -		
DR	Q62957	HSSP: P01130: 1LDL.		
DR	Q62957	InterPro: IPR000561; EGF-like.		
DR	Q62957	InterPro: IPR002172; LDL_recept_A.		
DR	Q62957	InterPro: IPR001862; MAC_perforin.		
DR	Q62957	InterPro: IPR000884; TSP1.		
DR	Q62957	Pfam: PF00057; ldl_recept_a; 1.		
DR	Q62957	Pfam: PF01823; MACPF; 1.		
DR	Q62957	Pfam: PF00090; TSP1; 1.		
DR	Q62957	PRINTS: PR00764; COMPLEMENTC9.		
DR	Q62957	SMART: SM00192; LDLa; 1.		
DR	Q62957	SMART: SM00457; MACPF; 1.		
DR	Q62957	SMART: SM00209; TSP1; 1.		
DR	Q62957	PROSITE: PS00022; EGF_1; UNKNOWN_1.		
DR	Q62957	PROSITE: PS01209; LDLRA_1; 1.		
DR	Q62957	PROSITE: PS0068; LDLRA_2; 1.		
DR	Q62957	PROSITE: PS00279; MAC_PERFORIN; 1.		
DR	Q62957	PROSITE: PS50092; TSP1; 1.		
DR	Q62957	Glycoprotein; Signal.		
FT	Q62957	SIGNAL 1 33		
FT	Q62957	POTENTIAL.		
FT	Q62957	COMPLEMENT COMPONENT C9.		
FT	Q62957	SEQUENCE 567 AA: 63762 MW: DBA9C728490AE68A CRC64;		



DT 01-JUN-1998 (TReMBLrel. 06, last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, last annotation update)  
DE OUTER MEMBRANE PROTEIN I PRECURSOR.  
GN POR.  
OS Neisseria gonorrhoeae.  
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=485;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=31493;  
RX MEDLINE=98258030; PubMed=9595666;  
RA Cooke S.J., Jolley K., Jison C.A., Young H., Heckels J.E.;  
RT "Naturally occurring isolates of Neisseria gonorrhoeae, which display  
RT anomalous serovar properties, express PIA/PIB hybrid porins, deletions  
RT in PIB or novel PIA molecules."  
RL FEMS Microbiol. Lett. 162:75-82(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=4001;  
RA Fudyk T.C., Maclean I.W., Simonsen J.N., Njagi E.N., Kimani J.,  
RA Brunham R.C., Plummer F.A.;  
RT "Genetic Diversity and Mosaicism at the por locus of Neisseria  
RT gonorrhoeae";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE  
CC (BY SIMILARITY).  
DR EMBL: AF015118; AAC38564.1; -;  
DR EMBL: AF090823; AAC79492.1; -;  
DR InterPro: IPR001702; Gram\_neg\_porin.  
DR Pfam: PF00267; Gram-ve\_porins; 1.  
DR PRINTS: PR00182; ECOLNEIPORIN.  
DR PROSITE: PS00576; GRAM\_NEG\_PORIN; 1.  
DR Outer membrane; Porin; Signal; Transmembrane.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 326 OUTER MEMBRANE PROTEIN I.  
SQ SEQUENCE 326 AA; 35577 MW; BF50F03EAF73BEE4 CRC64;

Query Match 35.9%; Score 52; DB 2; Length 326;  
Best Local Similarity 43.5%; Pred. No. 6.2;  
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 KCLGYHLDVSLAFSEISVGAFFN 23  
| | | | | : : : | | | : :  
Db 268 KGLVYHDLNDYDQVYVGAFFN 290

RESULT 6  
Q924L5 PRELIMINARY; PRT; 327 AA.  
AC Q924L5;  
DT 01-MAY-1999 (TReMBLrel. 10, Created)  
DT 01-MAY-1999 (TReMBLrel. 10, last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, last annotation update)  
DE OUTER MEMBRANE PROTEIN I PRECURSOR.  
GN POR.  
OS Neisseria gonorrhoeae.  
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=485;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=4266;  
RA Fudyk T.C., Maclean I.W., Simonsen J.N., Njagi E.N., Kimani J.,  
RA Brunham R.C., Plummer F.A.;  
RT "Genetic Diversity and Mosaicism at the por locus of Neisseria  
RT gonorrhoeae";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE  
CC (BY SIMILARITY).  
DR EMBL: AF090808; AAC79477.1; -;  
DR InterPro: IPR001702; Gram\_neg\_porin.

DR Pfam: PF00267; Gram-ve\_porins; 1.  
DR PRINTS: PR00182; ECOLNEIPORIN.  
DR PROSITE: PS00576; GRAM\_NEG\_PORIN; 1.  
KW Outer membrane; Porin; Signal; Transmembrane.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 327 OUTER MEMBRANE PROTEIN I.  
SQ SEQUENCE 327 AA; 35471 MW; FDB0D0FE8E53E32F CRC64;

Query Match 35.9%; Score 52; DB 2; Length 327;  
Best Local Similarity 43.5%; Pred. No. 6.3;  
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 KCLGYHLDVSLAFSEISVGAFFN 23  
| | | | | : : : | | | : :  
Db 269 KGLVYHDLNDYDQVYVGAFFN 291

RESULT 7  
O13834 PRELIMINARY; PRT; 3227 AA.  
ID O13834;  
AC O13834;  
DT 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-JAN-1998 (TReMBLrel. 05, last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, last annotation update)  
DE UBICUITIN SYSTEM PROTEIN.  
GN SPAC19D5.04.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomyces.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972H-;  
RA Devlin K., Churcher C.M.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972H-;  
RA Wood V., Bartell B.G., Rajandream M.A.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: Z99531; CAB16714.1; -;  
DR InterPro: IPR000569; HECT.  
DR InterPro: IPR001000; Glyco\_hydro\_10.  
DR InterPro: IPR001345; PG\_mutase.  
DR Pfam: PF00632; HECT; 1.  
DR PROSITE: PS00591; GLYCOSYL\_HYDROL\_F10; UNKNOWN\_1.  
DR PROSITE: PS50237; HECT; 1.  
DR PROSITE: PS00175; PG\_MUTASE; UNKNOWN\_1.  
DR SMART: SM00119; HECTC; 1.  
SQ SEQUENCE 3227 AA; 365031 MW; 07FC47AB79124575 CMC64;

Query Match 35.9%; Score 52; DB 3; Length 3227;  
Best Local Similarity 46.2%; Pred. No. 80;  
Matches 12; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 2 CIGYHLDVSLAFSEISVGAFFNKDDC 27  
| : | : | | | : | | | : | |  
Db 1339 CILYLLLEVLADSKKPDEFNFSEDC 1364

RESULT 8  
Q917L2 PRELIMINARY; PRT; 241 AA.  
ID Q917L2;  
AC Q917L2;  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, last annotation update)  
DE CG3066 PROTEIN.  
GN CG3066.  
OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 RN NCBI\_TaxID=7227;  
 RP  
 RC  
 RX  
 MEDLINE-20196006; PubMed-10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer E.G., Helt G., Nelson C.R., Pfeiffer B.D.,  
 RA April J.F., Agdayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bertana B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Harris N.L., Harvey D., Heiman T.J., Guan P., Harris M.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.R., Houck J.,  
 RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese H.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL  
 CC  
 CC -1- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.  
 DR EMBL: AE003676; AAC22127.1; -  
 DR FlyBase: FBgn0037515; CG3066.  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR001254; Trypsin.  
 DR Pfam: PF00089; trypsin.1.  
 DR PRINTS: PRO0722; CHYMOTRYPSIN.  
 DR SMART: SM00020; TRYP-Spc.1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE: PS00135; TRYPSIN\_SER.1.  
 KW Hydrolase; Serine protease.  
 SQ  
 SEQUENCE 241 AA; 27002 MW; AAA0228330F64C1E CRC64;

Query Match 32.4%; Score 47; DB 5; Length 241;  
 Best Local Similarity 40.0%; Pred. No 27;  
 Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

OY 8 DVSIAFSEISVGAEFNKDDC 27  
 Db 168 NIHLSQLCYGGFFRDSC 187

RESULT 9  
 ID 09SSX1  
 PRELIMINARY; PRT; 278 AA.

AC 09SSX1;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE OSEYAL.  
 GN OSEYAL.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriocaridaceae; Oryzaceae; Oryza.  
 RN NCBI\_TaxID=4530;  
 RP  
 RX  
 MEDLINE-99430869; PubMed-10503544;  
 RA Takeda Y., Hatanoe S., Sentoku N., Matsunaka M.;  
 RT "Homologs of animal eyes absent (eya) genes are found in higher  
 RL plants.";  
 RL Mol. Genet. 262:131-138(1999).  
 RL EMBL: AB028887; BAA85161.1;  
 SQ  
 SEQUENCE 278 AA; 31044 MW; 7BC1E2C41343B7A1 CRC64;

Query Match 32.4%; Score 47; DB 10; Length 278;  
 Best Local Similarity 40.7%; Pred. No 31;  
 Matches 11; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

OY 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27  
 Db 202 KCLLYRDLDDVAFENYSSWVGKLCQ 228

RESULT 10  
 ID 09SNG2  
 PRELIMINARY; PRT; 367 AA.

AC 09SNG2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE EST C72179(E153) CORRESPONDS TO A REGION OF THE PREDICTED GENE.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriocaridaceae; Oryzaceae; Oryza.  
 RN NCBI\_TaxID=4530;  
 RP  
 RX  
 SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC  
 RL clone: P0514G12.";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP000616; BAA85447.1; -  
 SQ  
 SEQUENCE 367 AA; 42367 MW; FF9F290DE82F273E CRC64;

Query Match 32.4%; Score 47; DB 10; Length 367;  
 Best Local Similarity 40.7%; Pred. No 42;  
 Matches 11; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

OY 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27  
 Db 235 KCLLYRDLDDVAFENYSSWVGKLCQ 261

RESULT 11  
 ID 09V322  
 PRELIMINARY; PRT; 391 AA.

AC 09V322;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE CG3066 PROTEIN.  
 GN CG3066.





SEQUENCE 203 AA; 22177 MW; 9D6E8BFC22DCBF73 CRC64;

## Query Match

Best Local Similarity 31.7%; Score 46; DB 12; Length 203;  
Matches 10; Conservative 4; Mismatches 6; Indels 6; Gaps 1;

QY 2 CLGHLDVSLAFSEISVGAFFNKDDC 27  
DB 143 CECGEJDIDL-----VGRQFSANDC 162

## RESULT 14

Q9XEFO

ID Q9XEFO

PRELIMINARY; PRT; 254 AA.

AC Q9XEFO

DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

GN HYPOTHEICAL 28.4 KDA PROTEIN.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicot; Rosidae;

OC NCBI\_Taxid=3702; Brassicales; Brassicaceae; Arabidopsids.

RX MEDLINE=99225673; Pubmed=10207155;

RT Wang M.L., Belmonte S., Kim U., Dolan M., Morris J.W., Goodman H.M.,

RT "A cluster of ABA-regulated genes on Arabidopsis thaliana BAC

RT 707M07."

RN Genome Res. 9:325-333(1999).

RP SEQUENCE FROM N.A.

RX MEDLINE=99225673; Pubmed=10207155;

RX MEDLINE=20083487; Pubmed=10617197;

RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,

RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,

RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,

RA Adams M.D., Shenn M., Vanden S.E., Umeyan L., Tallon L.J., Gill J.E.,

RA Copenhagen G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,

RA Salzberg S.L., Fraser C.M., Venter J.C.,

RT "Sequence and analysis of chromosome II of Arabidopsis thaliana."

RL Nature 402:761-768(1999).

DR EMBL; AF085279; AAP25935.1; -

DR EMBL; AC018721; AAP18734.1; -

DR InterPro; IPR001092; HLH\_dim.

DR Pfam; PF00010; HLH; 1.

DR SMART; SM00353; HLH; 1.

KW Hypothetical protein.

SQ SEQUENCE 254 AA; 28414 MW; EE3BCE353EE52619 CRC64;

## Query Match

Best Local Similarity 31.7%; Score 46; DB 10; Length 254;  
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 10 SLAFSEISVGAFFNKDDC 27  
DB 36 NLGFSFSSFGNFPADDC 53

## RESULT 15

Q67397

AC Q67397

PRELIMINARY; PRT; 292 AA.

DT 01-AUG-1998 (TREMBLrel. 07, Created)

DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)

DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

GN HYPOTHEICAL 33.2 KDA PROTEIN.

AC Q67397

OS Aquifex aeolicus.  
OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
OX NCBI\_Taxid=63363;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VF5;  
RX MEDLINE=98196666; Pubmed=9537320;  
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,  
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.,  
RT "The complete genome of the hyperthermophilic bacterium Aquifex  
aeolicus."

RL Nature 392:353-358(1998).

DR EMBL; AB000738; AAC07368.1; -

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 292 AA; 33244 MW; 38AB3FD436D35FEB CRC64;

## Query Match

Best Local Similarity 31.7%; Score 46; DB 2; Length 292;  
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 LGYHLDVSLAFSEISVG 19  
DB 261 LKYEFLRLAFSATLTG 277

Search completed: January 6, 2002, 09:10:38  
Job time: 451 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:01:52 ; Search time 42.93 Seconds  
(without alignments)  
46.587 Million cell updates/sec

Title: US-09-020-393b-18

Perfect score: 145  
Sequence: 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

A.Geneseq\_1101.\*  
1: /SIDS2/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	145	100.0	27	20	AAV27328
2	142	97.9	82	20	AAV27324
3	142	97.9	560	18	AAV18310
4	57	39.3	86	20	AAV27325
5	57	39.3	561	18	AAV18311
6	50	34.5	1263	22	AAV72637
7	44.5	30.7	162	21	AAV32113
8	44	30.3	582	22	AAU02929
9	44	30.3	662	21	AAU08632
10	44	30.3	1511	21	AAU08632
11	44	30.3	1544	21	AAU08632

12	44	30.3	1547	21	AAU08632	Arabidopsis thalia
13	43.5	30.0	330	22	AAU08632	Follicular conjunc
14	43	29.7	82	21	AAU08632	Human pancreatic c
15	43	29.7	82	22	AAU08632	Human colon cancer
16	43	29.7	90	21	AAU08632	Human ORF1446
17	43	29.7	154	21	AAU08632	Arabidopsis thalia
18	43	29.7	168	21	AAU08632	Arabidopsis thalia
19	43	29.7	213	20	AAU08632	Human secreted pro
20	43	29.7	271	21	AAU08632	Arabidopsis thalia
21	43	29.7	282	21	AAU08632	Arabidopsis thalia
22	43	29.7	282	21	AAU08632	Arabidopsis thalia
23	43	29.7	285	21	AAU08632	Arabidopsis thalia
24	43	29.7	306	21	AAU08632	Arabidopsis thalia
25	43	29.7	306	21	AAU08632	Arabidopsis thalia
26	43	29.7	309	21	AAU08632	Arabidopsis thalia
27	43	29.7	320	21	AAU08632	Arabidopsis thalia
28	43	29.7	320	21	AAU08632	Arabidopsis thalia
29	43	29.7	357	18	AAU08632	Prolyl-tRNA synthet
30	43	29.7	1070	21	AAU08632	S. epidermidis ope
31	42.5	29.3	434	22	AAU08632	Human P13 kinase p
32	42.5	29.3	591	16	AAU08632	Saccharomyces sp.
33	42	29.0	58	21	AAU08632	Arabidopsis thalia
34	42	29.0	74	21	AAU08632	Arabidopsis thalia
35	42	29.0	75	21	AAU08632	Arabidopsis thalia
36	42	29.0	77	22	AAU08632	Human colon cancer
37	42	29.0	117	19	AAU08632	Flag-amyloid prote
38	42	29.0	148	19	AAU08632	Open reading frame
39	42	29.0	298	22	AAU08632	C glutamic prote
40	42	29.0	305	22	AAU08632	S. epidermidis ope
41	42	29.0	484	17	AAU08632	Maize adenylosucc
42	42	29.0	484	17	AAU08632	Maize adenylosucc
43	42	29.0	484	20	AAU08632	Maize adenylosucc
44	42	29.0	516	21	AAU08632	Arabidopsis thalia
45	42	29.0	544	21	AAU08632	Arabidopsis thalia

#### ALIGNMENTS

RESULT 1	AAV27328	standard; peptide: 27 AA.
ID	AAV27328	
XX	AAV27328	
AC	AAV27328	
XX	05-NOV-1999	(first entry)
DF	XX	
XX	XX	
DE	Human C9 protein fragment (residues 359-384).	
XX	XX	
KW	CD59 mediated complement; human; Cd59 protein; mimetic;	
KW	tumor therapy; complement-mediated inflammation; immune disorder;	
KW	immunovascultitis; rheumatoid arthritis; scleroderma; C5b-9 complex.	
XX	XX	
OS	Homo sapiens.	
XX	XX	
PN	WO9440115-A2.	
XX	XX	
PD	12-AUG-1999.	
XX	XX	
PF	09-FEB-1999;	99WO-0502802.
XX	XX	
PR	09-FEB-1998;	98US-0020393.
XX	XX	
PA	(BLOO-) BLOOD CENT. RES. FOUND INC.	
PA	(OKLA-) OKLAHOMA MEDICAL RES. FOUND.	
XX	XX	
PI	Sims PJ;	
XX	XX	
DR	WPI; 1999-527301/44.	
XX	XX	
PT	Compounds modulating CD59 mediated complement activity, for	
XX	treatment of, e.g. Immunovascultitis	
XX	XX	

PS Example 2; Page 36; 75pp; English.  
XX  
CC The invention relates to compounds modulating CD59 mediated complement  
CC activity. It provides (i) molecules structurally mimicking human CD59  
CC amino acid residues 42-58 (region which serves as binding site for CD59  
CC C9 interactions) when they are in a spatial orientation which can  
CC inhibit the formation of the human C5b-9 complex. These mimetics  
CC specifically bind to amino acid residues 359-384 of human C9. (ii)  
CC molecules structurally mimicking C9 amino acids 359-384 when they are  
CC in a spatial orientation which can promote the formation of the C5b-9  
CC complex. Compounds that mimic CD59 can be used to increase CD59  
CC inhibition of C5b-9 complex assembly. This is especially useful in  
CC immune disorders and diseases such as immunovascularitis, rheumatoid  
CC arthritis, scleroderma. Compounds that mimic C9 can be used to promote  
CC C5b-9 complex assembly. This is useful in patients in need of complement  
CC activation. The composition can be administered as an adjunct to tumour  
CC therapy. The present sequence represents a human C9 protein fragment.  
XX  
S0 Sequence 27 AA;  
  
Query Match 100.0%; Score 145; DB 20; Length 27;  
Best Local Similarity 100.0%; Pred. No. 3 4e-17;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KCLGYHLDVSLAFSEISVGAFFNKDDC 27  
DB 1 KCLGYHLDVSLAFSEISVGAFFNKDDC 27  
|||||  
  
RESULT 2  
AAV27324  
ID AAV27324 standard; peptide; 82 AA.  
XX  
AC AAV27324;  
XX  
DT 05-NOV-1999 (first entry)  
XX  
DE Human C9 protein fragment (residues 334-415).  
XX  
KW CD59 mediated complement; human; CD59 protein; C9 protein; mimetic;  
KW tumour therapy; complement-mediated inflammation; immune disorder;  
KW immunovascularitis; rheumatoid arthritis; scleroderma; C5b-9 complex.  
OS Homo sapiens.  
XX  
PN MO9940115-A2.  
XX  
PD 12-AUG-1999.  
XX  
PF 09-FEB-1999; 99WO-US02802.  
XX  
PR 09-FEB-1998; 98US-0020393.  
XX  
PA (BLOO-) BLOOD CENT RES FOUND INC.  
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.  
XX  
PI Sims PJ;  
XX  
DR WPI; 1999-527301/44.  
XX  
PT Compounds modulating CD59 mediated complement activity, for  
PT treatment of, e.g. immunovascularitis  
XX  
PS Disclosure; Fig 10; 75pp; English.  
XX  
CC The invention relates to compounds modulating CD59 mediated complement  
CC activity. It provides (i) molecules structurally mimicking human CD59  
CC amino acid residues 42-58 (region which serves as binding site for CD59  
CC C9 interactions) when they are in a spatial orientation which can  
CC inhibit the formation of the human C5b-9 complex. These mimetics  
CC specifically bind to amino acid residues 359-384 of human C9. (ii)

CC molecules structurally mimicking C9 amino acids 359-384 when they are  
CC in a spatial orientation which can promote the formation of the C5b-9  
CC complex. Compounds that mimic CD59 can be used to increase CD59  
CC inhibition of C5b-9 complex assembly. This is especially useful in  
CC patients in need of suppression of complement-mediated inflammation, e.g.  
CC immune disorders and diseases such as immunovascularitis, rheumatoid  
CC arthritis, scleroderma. Compounds that mimic C9 can be used to promote  
CC C5b-9 complex assembly. This is useful in patients in need of complement  
CC activation. The composition can be administered as an adjunct to tumour  
CC therapy. The present sequence represents a human C9 protein fragment.  
XX  
S0 Sequence 82 AA;  
  
Query Match 97.9%; Score 142; DB 20; Length 82;  
Best Local Similarity 96.3%; Pred. No. 4 2e-16;  
Matches 26; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KCLGYHLDVSLAFSEISVGAFFNKDDC 27  
DB 25 KCLGYHLDVSLAFSEISVGAFFNKDDC 51  
:|||||  
  
RESULT 3  
AAW18310  
ID AAW18310 standard; Protein; 560 AA.  
XX  
AC AAW18310;  
XX  
DT 10-AUG-1997 (first entry)  
XX  
DE Human complement C9.  
XX  
KW Complement C9; CD59; C5b9 complex; tumour; therapy; contraceptive;  
KW antibody; inflammation.  
OS Homo sapiens.  
XX  
FH Key  
FT Binding-site 356..437  
FT /label= CD59 binding-site  
FT /note= "Claim 1 (amino acids 359-384)"  
XX  
FT Disulfide-bond 381..406  
XX  
PN WO9717987-A1.  
XX  
PD 22-MAY-1997.  
XX  
PF 08-NOV-1996; 96WO-US17940.  
XX  
PR 15-NOV-1995; 95US-0559492.  
XX  
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.  
XX  
PI Sims PJ;  
XX  
DR WPI; 1997-289058/26.  
DR N-PSDB; AAT68886.  
XX  
PT Modulating binding of CD59 to C9 complement component - uses agent  
PT that mimics or binds the C9-specific motif, used to activate  
PT complement in tumour therapy or to treat complement mediated  
PT inflammation  
XX  
PS Disclosure; Page 33-35; 51pp; English.  
XX  
CC Amino acid residues 359-384 of human complement C9 are critical  
CC for binding CD59 to C9, resulting in inhibition of C5b-9 complex  
CC assembly. Peptides that mimic this human C9-derived peptide  
CC sequence, and antibodies raised against such peptides, can be used  
CC to modulate binding of CD59 to C9. The peptides bind to endogenous  
CC CD59 to prevent the CD59 from inhibiting assembly of the C5b-9  
CC complex, thereby increasing complement-mediated activation of cells.

CC Inhibition of CD59 is useful as an adjuvant for tumour therapy and  
 CC as a contraceptive. The antibodies inhibit assembly of the C5b-9  
 CC complex by binding to C9. This is useful for disorders  
 CC characterised by excessive complement activation or complement-  
 CC mediated cytotoxicity.

XX Sequence 560 AA;

Query Match 97.9%; Score 142; DB 18; Length 560;

Best Local Similarity 96.3%; Pred. No. 4.5e-15;

Matches 26; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KCLGYHLDVSLAFSEISVGAEPFNKDDC 27

Db 380 rclqghldvslafseisvgaefnkdcc 406

RESULT 4

AAV27325 AAV27325 standard; peptide; 86 AA.

XX AAV27325;

DT 05-NOV-1999 (first entry)

XX Rabbit C9 protein fragment.

DE CD59 mediated complement; human; CD59 protein; C9 protein; mimetic;

KW tumour therapy; complement-mediated inflammation; immune disorder;

KW immunovascularitis; rheumatoid arthritis; scleroderma; C5b-9 complex;

XX rabbit.

OS Oryctolagus sp.

XX WO9940115-A2.

PN 12-AUG-1999.

PD 09-FEB-1999; 99MO-US02802.

PF 09-FEB-1998; 98US-0020393.

PR (BLOO-) BLOOD CENT RES FOUND INC.

PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.

XX Sims PJ;

PI WPI; 1999-527301/44.

DR Compounds modulating CD59 mediated complement activity, for

XX treatment of, e.g. immunovascularitis

PT Disclosure; Fig 10; 75pp; English.

XX The invention relates to compounds modulating CD59 mediated complement

CC activity. It provides (1) molecules structurally mimicking human CD59

CC amino acid residues 42-58 (region which serves as binding site for C5b-9

CC -C9 interactions) when they are in a spatial orientation which can

CC inhibit the formation of the human C5b-9 complex. These mimetics

CC specifically bind to amino acid residues 359-384 of human C9. (11)

CC molecules structurally mimicking C9 amino acids 359-384 when they are

CC in a spatial orientation which can promote the formation of the C5b-9

CC complex. Compounds that mimic CD59 can be used to increase CD59

CC inhibition of C5b-9 complex assembly. This is especially useful in

CC patients in need of suppression of complement-mediated inflammation, e.g.

CC immune disorders and diseases such as immunovascularitis, rheumatoid

CC arthritis, scleroderma. Compounds that mimic C9 can be used to promote

CC C5b-9 complex assembly. This is useful in patients in need of complement

CC activation. The composition can be administered as an adjunct to tumour

CC therapy. The present sequence represents a rabbit C9 protein fragment.

XX Sequence 86 AA;

Query Match 39.3%; Score 57; DB 20; Length 86;  
 Best Local Similarity 38.7%; Pred. No. 0.063;  
 Matches 12; Conservative 8; Mismatches 7; Indels 4; Gaps 1;

QY 1 KCLGYHLDVSLAFSEISVGAEPFNKDDC 27

Db 25 kclgfdldslinipksaglsltgqanknc 55

RESULT 5

AAW18311 AAW18311 standard; Protein; 561 AA.

XX AAW18311;

DT 10-AUG-1997 (first entry)

XX Rabbit complement C9.

DE Complement C9; CD59; C5b9 complex; tumour; therapy; contraceptive;

KW antibody; inflammation.

OS Oryctolagus cuniculus.

XX Key Location/Qualifiers

FT Disulfide-bond 389..418

PD WO9717987-A1.

PN 22-MAY-1997.

PE 08-NOV-1996; 96MO-US17940.

PR 15-NOV-1995; 95US-0559492.

XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.

PA Sims PJ;

PI WPI; 1997-289058/26.

DR N-PSDB; T688887.

XX Modulating binding of CD59 to C9 complement component - uses agent

PT that mimics or binds the C9-specific motif, used to activate

PT complement in tumour therapy or to treat complement mediated

PT inflammation

XX Disclosure; Page 36-38; 51pp; English.

XX A polypeptide (AAW18311) comprises rabbit complement C9. Chimeric

CC proteins were constructed in which the segment of C9 corresponding

CC to the putative CD59 binding site were interchanged between rabbit

CC and human C9 (see also AAW18310). These were tested for haemolytic

CC activity and for sensitivity to inhibition by membrane CD59.

CC Substr. of human C9 residues 334-415 into rabbit C9 resulted in a

CC protein that was indistinguishable from human C9 in its sensitivity

CC to inhibition by CD59. When this same segment of human C9 was

CC replaced by the corresponding rabbit sequence (aa339-424), the

CC chimera was indistinguishable from rabbit C9 and was virtually

CC unaffected by the presence of membrane CD59.

XX Sequence 561 AA;

Query Match 39.3%; Score 57; DB 18; Length 561;

Best Local Similarity 38.7%; Pred. No. 0.63;

Matches 12; Conservative 8; Mismatches 7; Indels 4; Gaps 1;

QY 1 KCLGYHLDVSLAFSEISVGAEPFNKDDC 27

Db 388 kclgfdldslinipksaglsltgqanknc 418

## RESULT 6

AAV72637  
ID AAV72637 standard; Protein; 1263 AA.  
XX  
XX AAV72637;  
AC  
XX  
DT 02-MAY-2001 (first entry)  
XX  
DE Exophiala splinifera p-glycoprotein, a fumonisin degradative enzyme.  
XX  
KW Fumonisin; degradative enzyme; transport enzyme; flavin monooxygenase;  
KW aldehyde dehydrogenase; perasease; p-glycoprotein; fumonisin esterase;  
KW detoxification; mycotoxin; animal feed; human feed; silage;  
KW transgenic plant; transgenic animal; microbial spray.  
XX  
OS Exophiala splinifera.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 157  
FT /label= Unknown  
FT /note= "Encoded by CAR"  
XX  
PN W0200105980-A1.  
XX  
PD 25-JAN-2001.  
XX  
PF 14-JUL-1999; 99WO-US15824.  
XX  
PR 14-JUL-1999; 99WO-US15824.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
XX (CURA-) CURAGEN CORP.  
XX  
PI Duvick JP, Maddox J, Gilliam J, Folkerts O, Crasta OR;  
XX WPI; 2001-147345/15.  
DR N-PSDB; AAD02694, AAD02695.  
XX  
PT Novel polynucleotides encoding Exophiala degradative or transport  
XX enzyme which is useful for detoxifying fumonisin or structurally  
XX related mycotoxin during processing of grain for human or animal food  
XX consumption -  
XX  
PS Claim 1c; Page 77-82; 90pp; English.  
XX  
CC The patent discloses novel polynucleotides encoding Exophiala splinifera  
CC fumonisin degradative or transport enzymes such as flavin monooxygenase,  
CC aldehyde dehydrogenase, perasease and p-glycoprotein. A fumonisin esterase  
CC enzyme with at least one fumonisin degradative enzyme is useful for  
CC detoxifying fumonisin or a structurally related mycotoxin present in  
CC harvested grain, on application to a plant or to harvested grain  
CC during processing, or to processed grain that is to be used as animal  
CC or human feed, or as silage. Nucleotide fragments of the present  
CC invention are useful as probes and primers. They can be introduced  
CC into microorganisms that multiply on plants to deliver enzymes to  
CC potential target crops. The genes encoding the degrading enzymes are  
CC introduced via a vector into a microbial host and the transformed host  
CC is supplied to the environment, plants or animals for reducing the  
CC pathogenicity of a fungus producing fumonisin. The genes of the  
CC invention are fermented in a bacterial host and the resulting bacteria  
CC can be used alone or in combination to engineer microbes or other  
CC organisms to metabolise fumonisin and resist its toxic effects.  
CC The present protein sequence is p-glycoprotein, a fumonisin  
CC degradative enzyme from Exophiala splinifera.  
XX  
SQ Sequence 1263 AA:

Query Match 34.5%; Score 50; DB 22; Length 1263;  
Best Local Similarity 42.1%; Pred. No. 25;

Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;  
QY 1 KCLGYHLDVSLAFSEISVG 19  
:|:|:| |:|:|:  
Db 922 rcfgfhlsqsmeflatalg 940

## RESULT 7

AAV32113  
ID AAV32113 standard; Protein; 162 AA.  
XX  
XX AAV32113;  
AC  
XX  
DT 01-FEB-2000 (first entry)  
XX  
DE Mouse 15 kDa selenoprotein.  
XX  
KW Selenoprotein; selenium; differential expression; tumour; mouse;  
KW prostate cancer; diagnosis; polymorphism.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 93  
FT /label= OTHER  
FT /note= "selenocysteine"  
XX  
PN W09951637-A1.  
XX  
PD 14-OCT-1999.  
XX  
PF 06-APR-1999; 99WO-US07560.  
XX  
PR 06-APR-1998; 98US-0080850.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX Gladyshev VN, Wootton JC, Hatfield DL, Jeang K;  
XX WPI; 2000-013045/01.  
DR N-PSDB; AA34469.  
XX  
PT New isolated selenoprotein polypeptides, used to develop products for  
XX detecting susceptibility to or treating cancers e.g. prostate cancers  
XX -  
XX  
PS Claim 1; Page 60; 67pp; English.  
XX  
CC The present sequence represents a mouse 15 kDa selenoprotein that  
CC is differentially expressed in cancer cells, such as cancer cells.  
CC It includes a selenocysteine residue at position 93 that is encoded  
CC by a TGA codon in the cDNA (see AA34468). Polymorphisms in the  
CC human 15 kDa selenoprotein gene are associated with susceptibility  
CC to cancer. Claimed methods of detecting the presence, or of  
CC quantifying the level of expression, of the 15 kDa selenoprotein  
CC involve contacting a sample with a specific binding agent, such as  
CC a polyclonal antibody or monoclonal antibody, that specifically  
CC binds to the 15 kDa selenoprotein. A claimed method for dietary  
CC regulation comprises detecting an abnormally low expression of a  
CC mammalian 15 kDa selenoprotein in a subject and, if the level is  
CC below normal, enhancing the level by providing additional selenium  
CC in the diet. The susceptibility to cancer of a subject having an  
CC increased predetermined genetic susceptibility to cancer is reduced  
CC by administering a mammalian 15 kDa selenoprotein or by  
CC overexpressing the selenoprotein using gene therapy.  
XX  
SQ Sequence 162 AA:

Query Match 30.7%; Score 44.5; DB 21; Length 162;  
Best Local Similarity 42.3%; Pred. No. 16;  
Matches 11; Conservative 5; Mismatches 9; Indels 1; Gaps 1;







PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161933.  
PR 29-OCT-1999; 99US-0162142.  
PR 29-OCT-1999; 99US-0162143.

Query Match 30.38; Score 44; DB 21; Length 1511;  
Best Local Similarity 47.68; Pred. No. 3.1e+02;  
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 6 HLDVSLAFSEISVCAFEFKKD 26  
||||||| : :| : :|  
Db 906 hldvslavawamqilfred 926

## RESULT 11

AAAG38527  
ID AAAG38527 standard; Protein; 1544 AA.

AC AAG38527;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 47542.

KW Protein identification; signal transduction pathway; metabolic pathway;

KM hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

OS Arabidopsis thaliana.

XX Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.  
PD 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137282.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.

PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158363.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159337.
PR	14-OCT-1999;	99US-0159633.
PR	18-OCT-1999;	99US-0159638.
PR	21-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160776.
PR	21-OCT-1999;	99US-0160815.
PR	21-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161820.
PR	28-OCT-1999;	99US-0161992.
PR	29-OCT-1999;	99US-0162142.

```

Query Match          30.3%;   Score 44;   DB 21;   Length 1544;
Best Local Similarity 47.6%;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY      6 HLDVSLAFSEISVGAEPNKDD 26
      ||||| : : | : |
Db      939 hldvslavawafmgqifresd 959

RESULT 12
AAG38526
ID      AAG38526 standard; Protein; 1547 AA.
XX
XX      AAG38526;
AC
XX
XX      18-OCT-2000 (first entry)
DT
XX
DE      Arabidopsis thaliana protein fragment SPQ ID NO: 47541.
XX
XX      Protein identification; signal transduction pathway; metabolic pathway;
KW      hybridisation assay; genetic mapping; gene expression control; promoter;
XX      termination sequence.
KW
XX      Arabidopsis thaliana.
OS
XX      EP1033405-A2.
PN
XX
XX      06-SEP-2000.
PD
XX
XX      25-FEB-2000; 2000EP-0301439.
PF
XX
XX      25-FEB-1999; 99US-0121825.
PR      05-MAR-1999; 99US-0123180.

```

PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 04-MAY-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.

XX	RESULT 13
XX	ID AAG62569
XX	AAG62569 standard; Protein; 330 AA.
XX	AC AAG62569;
XX	DT 05-SEP-2001 (first entry)
DE	Follicular conjunctivitis related adenoviral protein SEQ ID NO: 10.
XX	XX Follicular conjunctivitis; antiserum; antiviral; vaccine; infection
XX	OS Mastadenovirus.
PN	JP2001095583-A.
XX	PD 10-APR-2001.
XX	PD 30-SEP-1999; 99JP-0278661.
XX	PR 30-SEP-1999; 99JP-0278661.
PA	(ITON/) ITO N.
XX	DR WPI; 2001-341249/36.
XX	DR N-PSDB; AAAH42033.
XX	PT New adenovirus for the prevention and treatment of Ad infection -
XX	Disclosure; Page 23-25; 45pp; Japanese.
XX	

nm0207173 to AAC99231 encode the human pancreatic cancer associated proteins, called pancreatic cancer antigens, given in AAB54466 to AAB54466. The human pancreatic cancer antigens have cytostatic, neuroprotective, neurotropic, immunomodulatory, relaxant, contraceptive, gynaecological, cardiant and antiinflammatory activities, and can be used in gene therapy. The polynucleotide and proteins can be used for preventing, treating, or ameliorating a medical condition or in assays subject. Binding partners to the proteins and the activity of the proteins can be identified. The pancreatic cancer antigens can be used to detect, treat or prevent pancreatic disorders, especially cancer. Agonists and antagonists to the antigens can be screened for. The

CC pancreatic cancer antigen polynucleotides can be used to design nucleic  
CC acid hybridisation probes that can be used in chromosome mapping, linkage  
CC analysis, tissue identification and/or typing and a variety of forensic  
CC and diagnostic methods. The proteins can be used to generate antibodies  
CC which are used to purify, detect and target the polypeptides, including  
CC both in vivo and in vitro diagnostic and therapeutic methods. The  
CC proteins can be used to treat or prevent neural, immune system, muscular,  
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or  
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent  
CC sequences used in the exemplification of the present invention.

XX  
SQ Sequence 82 AA:

Query Match 29.7%; Score 43; DB 21; Length 82;  
Best Local Similarity 33.3%; Pred. No. 13;  
Matches 8; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 2 CIGYHLDVSLAFSEISVGAEFNKD 25  
| | | | | : | : | : |  
Db 4 cfayhrdislmavppcrvayqtdvd 27

RESULT 15  
AAG74079  
ID AAG74079 standard; Protein; 82 AA.  
XX  
AC AAG74079;  
XX  
DT 03-SEP-2001 (first entry)  
XX  
DE Human colon cancer antigen protein SEQ ID NO:4843.  
XX  
KM Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KM colorectal carcinoma; chromosome 11.  
XX  
OS Homo sapiens.  
XX  
PN WC200122920-A2.  
XX  
PD 05-APR-2001.  
XX  
PF 28-SEP-2000; 2000MO-US26524.  
XX  
PR 29-SEP-1999; 99US-0157137.  
PR 03-NOV-1999; 99US-0163280.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Ruben SM, Barash SC, Birse CE, Rosen CA;  
XX  
DR WPI: 2001-235357/24.  
DR N-PSDB; AAH33510.  
XX  
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
PT useful for preventing, diagnosing and/or treating colorectal cancers -  
XX  
PS Claim 11: Page 6627-6628; 9803pp; English.  
XX  
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
CC the proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene  
CC therapy and vaccine production. N and P may be used in the prevention,  
CC diagnosis and treatment of diseases associated with inappropriate P  
CC expression. For example, N and P may be used to treat disorders  
CC associated with decreased expression by rectifying mutations or deletions  
CC in a patient's genome that affect the activity of P by expressing  
CC inactive proteins or to supplement the patients own production of P.  
CC Additionally, N may be used to produce the colon cancer-associated Ps,  
CC by inserting the nucleic acids into a host cell and culturing the cell  
CC to express the proteins. N and P can be used in the prevention, diagnosis  
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204

CC and AAB77789 represent sequences used in the exemplification of the  
CC present invention.  
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
CC missing at time of publication, meaning no sequences are present for  
CC SEQ ID NO:1027 to 1052, 7921 and 7922.  
XX  
SQ Sequence 82 AA:

Query Match 29.7%; Score 43; DB 22; Length 82;  
Best Local Similarity 33.3%; Pred. No. 13;  
Matches 8; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 2 CIGYHLDVSLAFSEISVGAEFNKD 25  
| | | | | : | : | : |  
Db 4 cfayhrdislmavppcrvayqtdvd 27

Search completed: January 6, 2002, 09:04:12  
Job time: 140 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:01:52 ; Search time 25.89 Seconds  
(without alignments)  
23.468 Million cell updates/sec

Title: US-09-020-393b-18

Perfect score: 145  
Sequence: 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCRTUS.COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/Backfillset1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	142	97.9	82	2	US-08-559-492-3
2	142	97.9	557	1	US-08-313-288B-16
3	142	97.9	560	2	US-08-559-492-5
4	57	39.3	86	2	US-08-559-492-4
5	57	39.3	561	2	US-08-559-492-12
6	42	29.0	117	2	US-08-729-345-3
7	42	29.0	148	1	US-08-565-386-8
8	42	29.0	359	4	US-09-347-798-12
9	42	29.0	484	1	US-08-361-611-4
10	42	29.0	484	1	US-08-565-655-4
11	42	29.0	484	2	US-08-946-967-4
12	41.5	28.6	365	3	US-09-231-529-3
13	41.5	28.6	365	4	US-08-977-816-3
14	41.5	28.6	619	3	US-08-262-220-6
15	41.5	28.6	619	3	US-08-471-733-6
16	41.5	28.6	619	3	US-08-468-878-6
17	41.5	28.6	619	4	US-08-750-494-6
18	41.5	28.6	990	2	US-08-392-625-20
19	41.5	28.6	990	2	US-08-466-961A-20
20	41.5	28.6	990	2	US-08-645-193B-15
21	41	28.3	211	1	US-08-631-607-4
22	41	28.3	560	2	US-08-948-569A-10
23	41	28.3	560	2	US-09-188-469-10
24	41	28.3	564	4	US-09-397-238A-10
25	41	28.3	564	2	US-08-948-569A-12
26	41	28.3	564	2	US-09-188-469-12
27	41	28.3	564	4	US-09-397-238A-12

28	41	28.3	2647	2	US-08-583-562B-8	Sequence 8, App1
29	41	28.3	2647	2	US-08-779-113-8	Sequence 8, App1
30	40	27.6	186	2	US-08-960-022-18	Sequence 18, App1
31	40	27.6	393	4	US-09-347-798-18	Sequence 18, App1
32	40	27.6	721	4	US-09-146-221-4	Sequence 4, App1
33	40	27.6	724	4	US-09-146-221-6	Sequence 6, App1
34	40	27.6	727	4	US-09-146-221-8	Sequence 8, App1
35	40	27.6	971	3	US-09-112-450-2	Sequence 2, App1
36	40	27.6	2471	3	US-09-112-450-4	Sequence 4, App1
37	39.5	27.2	621	3	US-08-262-220-8	Sequence 8, App1
38	39.5	27.2	621	3	US-08-471-733-8	Sequence 8, App1
39	39.5	27.2	621	3	US-08-468-878-8	Sequence 8, App1
40	39.5	27.2	621	4	US-08-750-494-8	Sequence 8, App1
41	39	26.9	282	3	US-08-733-230-6	Sequence 6, App1
42	39	26.9	282	4	US-08-953-326-6	Sequence 6, App1
43	39	26.9	529	1	US-08-152-019A-40	Sequence 40, App1
44	39	26.9	529	3	US-08-650-599A-3	Sequence 3, App1
45	39	26.9	603	3	US-08-482-677-8	Sequence 8, App1

#### ALIGNMENTS

```
RESULT 1
US-08-559-492-3
; Sequence 3, Application US/08559492
; Patent No. 5843884
;
GENERAL INFORMATION:
; APPLICANT: Sims, Peter J.
; TITLE OF INVENTION: C9 Complement Inhibitor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1201 West Peachtree
; STREET: Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/559,492
; FILING DATE: 15-NOV-1995
;
CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMR154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-873-8794
; TELEFAX: 404-873-8795
;
INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 82 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
US-08-559-492-3
```

Query Match 97.9%; Score 142; DB 2; Length 82;  
Best local similarity 96.3%; Pred. No. 1.8e-16;  
Matches 26; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27  
DB 25 KCLGYHLDVSLAFSEISVGAEFNKDDC 51

```
RESULT 2
US-08-313-288B-16
; Sequence 16, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; NUMBER OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-313-288B-16

Query Match          97.9%; Score 142; DB 1; Length 557;
Best Local Similarity 96.3%; Pred. No. 2.1e-15;
Matches 26; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KCLGYHLDVSLAFSEISVGAEPFNKDDC 27
Db 377 RCLGYHLDVSLAFSEISVGAEPFNKDDC 403

RESULT 3
US-08-559-492-5
; Sequence 5, Application US/08559492
; Patent No. 5843884
; GENERAL INFORMATION:
; APPLICANT: Sims, Peter J.
; TITLE OF INVENTION: C9 Complement Inhibitor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1201 West Peachtree
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; US-08-559-492-5
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/559,492
; FILING DATE: 15-NOV-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-873-8794
; TELEFAX: 404-873-8795
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-559-492-5
```

```
Query Match          97.9%; Score 142; DB 2; Length 560;
Best Local Similarity 96.3%; Pred. No. 2.1e-15;
Matches 26; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KCLGYHLDVSLAFSEISVGAEPFNKDDC 27
Db 380 RCLGYHLDVSLAFSEISVGAEPFNKDDC 406
```

```
RESULT 4
US-08-559-492-4
; Sequence 4, Application US/08559492
; Patent No. 5843884
; GENERAL INFORMATION:
; APPLICANT: Sims, Peter J.
; TITLE OF INVENTION: C9 Complement Inhibitor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1201 West Peachtree
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/559,492
; FILING DATE: 15-NOV-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-873-8794
; TELEFAX: 404-873-8794
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-559-492-4
```



Query Match 39.3%; Score 57; DB 2; Length 86;  
Best Local Similarity 38.7%; Pred. No. 0.025;  
Matches 12; Conservative 8; Mismatches 7; Indels 4; Gaps 1;

QY 1 KCLGYHLDVSLAF----SEISVGAEPFNKDDC 27  
DB 25 KCLGFDLDSLNPGRKAGLSLTGQANKNNC 55

RESULT 5  
US-08-559-492-12  
; Sequence 12, Application US/08559492  
; Patent No. 5843884  
; GENERAL INFORMATION:  
; APPLICANT: Sims, Peter J.  
; TITLE OF INVENTION: C9 Complement Inhibitor  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patrea L. Pabst  
; STREET: 2800 One Atlantic Center, 1201 West Peachtree  
; STREET: Street  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30309-3450  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/559,492  
; FILING DATE: 15-NOV-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: OMRF154  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404-873-8794  
; TELEFAX: 404-873-8795  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 561 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-559-492-12

Query Match 39.3%; Score 57; DB 2; Length 561;  
Best Local Similarity 38.7%; Pred. No. 0.27;  
Matches 12; Conservative 8; Mismatches 7; Indels 4; Gaps 1;

QY 1 KCLGYHLDVSLAF----SEISVGAEPFNKDDC 27  
DB 388 KCLGFDLDSLNPGRKAGLSLTGQANKNNC 418

RESULT 6  
US-08-729-345-3  
; Sequence 3, Application US/08729345  
; Patent No. 5849999  
; GENERAL INFORMATION:  
; APPLICANT: Neve, Rachael L.  
; APPLICANT: Berger-Sweeney, Joanne  
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMAL  
; TITLE OF INVENTION: EXPRESSING FLAG-APP-C100 FUSION PROTEIN  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/729,345  
; FILING DATE: 16-OCF-1996  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 04843/027001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 117 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
US-08-729-345-3

Query Match 29.0%; Score 42; DB 2; Length 117;  
Best Local Similarity 44.4%; Pred. No. 11;  
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 8 DVSLAFSEISVGAEPFNKDD 25  
DB 8 DKSMGISEVKMDAEFRHD 25

RESULT 7  
US-08-565-386-8  
; Sequence 8, Application US/08565386  
; Patent No. 5741697  
; GENERAL INFORMATION:  
; APPLICANT: Bayoill, Patrick M.  
; APPLICANT: Hsia, Ru-ching  
; TITLE OF INVENTION: BACTERIOPHAGE OF CHLAMYDIA PSITTACI  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Roches  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/565,386  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Timian, Susan J.  
; REGISTRATION NUMBER: 34,103  
; REFERENCE/DOCKET NUMBER: 176/60040  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 716-263-1636  
TELEFAX: 716-263-1600  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 148 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-565-386-8

Query Match 29.0%; Score 42; DB 1; Length 148;  
Best Local Similarity 36.8%; Pred. No. 15;  
Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

OY 9 LVSLAFSEISVGAEFNKDDC 27  
12 VMKFTQKSLTQOHNKDEC 30

RESULT 8  
US-09-347-798-12  
Sequence 12, Application US/09347798  
Patent No. 6242256  
GENERAL INFORMATION:  
APPLICANT: Cahoon, Edgar B.  
APPLICANT: Cahoon, Rebecca E.  
APPLICANT: Hitz, William D.  
APPLICANT: Rafalski, Antoni  
TITLE OF INVENTION: Ornithine Biosynthesis Enzymes  
FILE REFERENCE: BB-1174-B  
CURRENT APPLICATION NUMBER: US/09/347,798  
CURRENT FILING DATE: 1999-07-02  
EARLIER APPLICATION NUMBER: 60/093,209  
EARLIER FILING DATE: July 17, 1998  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 12  
LENGTH: 359  
TYPE: PRT  
ORGANISM: zea mays  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (289)  
US-09-347-798-12

Query Match 29.0%; Score 42; DB 4; Length 359;  
Best Local Similarity 33.3%; Pred. No. 48;  
Matches 7; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

OY 7 LDVSLAFSEISVGAEFNKDDC 27  
Db 267 MSVSRSFQITVDGDTSTNDC 287

RESULT 9  
US-08-361-611-4  
Sequence 4, Application US/08361611  
Patent No. 5519125  
GENERAL INFORMATION:  
APPLICANT: Potter, Sharon L.  
APPLICANT: Ward, Eric R.  
TITLE OF INVENTION: Plant Adenylosuccinate Synthetase and  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA

ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/361,611  
FILING DATE:  
CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919) 541-8614  
TELEFAX: (919) 541-8689  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-361-611-4

Query Match 29.0%; Score 42; DB 1; Length 484;  
Best Local Similarity 45.0%; Pred. No. 70;  
Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

OY 7 LDVSLAFSEISVGAEFNKDD 26  
Db 391 LDVSLGSEIKVGYSTQTD 410

RESULT 10  
US-08-565-655-4  
Sequence 4, Application US/08565655  
Patent No. 5688939  
GENERAL INFORMATION:  
APPLICANT: Potter, Sharon L.  
APPLICANT: Ward, Eric R.  
TITLE OF INVENTION: Plant Adenylosuccinate Synthetase and  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ciba Patent Department  
STREET: 540 White Plains Rd., POB 2005  
CITY: Tarrytown  
STATE: NY  
COUNTRY: USA  
ZIP: 10591-9005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/565,655  
FILING DATE:  
CLASSIFICATION: 210  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/361,611  
FILING DATE: 12-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919) 541-8614  
TELEFAX: (919) 541-8689  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-565-655-4

Query Match 29.0%; Score 42; DB 1; Length 484;  
Best Local Similarity 45.0%; Pred. No. 70;  
Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 7 LDVSLAFSEISVGAENKDD 26  
||| ||| ||| : : :  
Db 391 LDVSLGSLSEIKVGVSTQTD 410

RESULT 11  
US-08-946-967-4  
Sequence 4, Application US/08946967  
Patent No. 5882869

GENERAL INFORMATION:  
APPLICANT: Potter, Sharon L  
APPLICANT: Ward, Eric R  
TITLE OF INVENTION: Plant Adenylosuccinate Synthetase and  
TITLE OF INVENTION: DNA Coding Therefor  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/946,967  
FILING DATE:  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919) 541-8614  
TELEFAX: (919) 541-8689  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-946-967-4

Query Match 29.0%; Score 42; DB 2; Length 484;  
Best Local Similarity 45.0%; Pred. No. 70;  
Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 7 LDVSLAFSEISVGAENKDD 26  
||| ||| ||| : : :  
Db 391 LDVSLGSLSEIKVGVSTQTD 410

RESULT 12  
US-09-231-529-3  
Sequence 3, Application US/09231529  
Patent No. 6096308

GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Corley, Neil C.

APPLICANT: Shah, Purvi  
TITLE OF INVENTION: HUMAN PROTEIN KINASE AND KINASE INHIBITORS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/231,529  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/977,816  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0429 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 365 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: KIDNOT25  
CLONE: 353694  
US-09-231-529-3

Query Match 28.6%; Score 41.5; DB 3; Length 365;  
Best Local Similarity 38.5%; Pred. No. 59;  
Matches 10; Conservative 4; Mismatches 3; Indels 9; Gaps 1;

QY 1 KCLG-----YHLDVSLAFSEIS 17  
||:| :||: ||||  
Db 69 KCVGKEDTGSMTMMDVACAFDEIS 94

RESULT 13  
US-08-977-816-3  
Sequence 3, Application US/08977816  
Patent No. 6194186

GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Corley, Neil C.  
APPLICANT: Shah, Purvi  
TITLE OF INVENTION: HUMAN PROTEIN KINASE AND KINASE INHIBITORS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS

```

Query March          28.6%  Score 41.5;  DB 4;  Length 365;
Best Local Similarity 38.5%  Pred. No. 59;
Matches 10;  Conservative 4;  Mismatches 3;  Indels 9;  Gaps 1
QY      1  KCLG-----YHDSVLSARESEIS 17
          ||:|  ||:|  ||  |||
Db      69  KCVGKPEDTGSMTATMNDVACAFDEIS 94

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	28.6%	Score 41.5;	DB 3;	Length 619;
Best Local Similarity	45.0%;	Pred. No. 1.2e+02;		
Matches	9;	Conservative	5;	Mismatches 5;
				Indels 1;
				Gaps 1

```
QY      8  DVSLESE-ISVGAEPNKDD 26
          :: : |||  | : | : |||
Db      372 ELGIAFSTGASIGLANNKDD 391
```

RESULT 15  
US-08-471-733-6  
; Sequence 6, Application US/08471733  
; Patent No. 6068842

APPLICANT: BERGSTROM SVEN  
APPLICANT: BARBOUR ALAN G.  
TITLE OF INVENTION: NEM 66 KDA ANTIGEN FROM BORRELLIA  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 SEVENTH STREET, N.W.  
CITY: WASHINGTON  
COUNTRY: USA  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,733  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

```

1 APPLICATION NUMBER: 08/262,220
2 FILING DATE: 20-JUN-1994
3 ATTORNEY/AGENT INFORMATION:
4 NAME: COOPER, IVER P.
5 REGISTRATION NUMBER: 28,005
6 REFERENCE/DOCKET NUMBER: BERGSTROM-3
7 TELECOMMUNICATION INFORMATION:
8 TELEPHONE: 202-628-5197
9 TELEFAX: 202-737-3528
10
11 TELEX: 248633
12
13 INFORMATION FOR SEQ ID NO: 6:
14 SEQUENCE CHARACTERISTICS:
15 LENGTH: 619 amino acids
16 TYPE: amino acid
17 TOPOLOGY: linear
18
19 MOLECULE TYPE: protein
20
21 US-08-471-733-6

```

Query March	28.6%	Score 41.5;	DB 3;	Length 619;
Best Local	Similarity 45.0%;	Pred. No. 1.2e+02;		
Matches 9;	Conservative 5;	Mismatches 5;	Indels 1;	Gaps 1.
Oy	8 DVSIAFSE-ISVGAERNKDD 26			
	:::     :::			
Db	372 ELGIARSTGASIGLAWNKDD 391			

Search completed: January 6, 2002, 09:05:06  
Job time: 194 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:23:06 ; Search time 12.65 seconds  
(without alignments)  
102.369 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58

Perfect score: 96  
Sequence: 1 FEHCNFDVTTRELENE 17

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 2975

Minimum DB seq length: 0  
Maximum DB seq length: 17

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database : PIR\_68:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	29.2	15	2 A61612	allatostatin - tob
2	25	26.0	15	2 PC2215	fibrinogenolytic p
3	24	25.0	13	2 H56046	urinary tract ston
4	23	24.0	9	2 A11497	transaldolase (EC
5	23	24.0	14	2 PS0371	hypothetical prote
6	23	24.0	17	2 A61117	somatostatin precu
7	22	22.9	16	2 A59042	alpha-conotoxin Ep
8	22	22.9	16	2 G49039	T-cell receptor be
9	21	21.9	12	2 T46794	hypothetical prote
10	21	21.9	13	2 S08575	botulinum neurotox
11	21	21.9	16	2 I52226	aldehyde dehydroge
12	21	21.9	8	2 S71919	transaldolase (EC
13	20	20.8	9	2 A12872	proteohondolisin c
14	20	20.8	14	2 PT0077	GTP-binding protel
15	20	20.8	14	2 S29486	nitrogenase cofact
16	20	20.8	15	2 S32677	seed storage prote
17	20	20.8	15	2 PA0009	bombolitin IV - Am
18	19	19.8	17	2 D22595	Ig heavy chain CRD
19	19	19.8	9	2 PT0272	peptidylglycine mo
20	19	19.8	9	2 A42266	T-cell receptor be
21	19	19.8	10	2 PH0895	aminotransferase c
22	19	19.8	11	4 PC2124	Ig H chain V-D-J r
23	19	19.8	14	2 PH1627	ribosomal protein
24	19	19.8	15	2 S36889	H+-transsporting AT
25	19	19.8	15	2 PT0095	photosystem II chl
26	19	19.8	17	2 S20490	hypothetical prote
27	18	18.8	7	2 S08606	hypothetical prote
28	18	18.8	10	2 S27873	Ig heavy chain CRD
29	18	18.8	10	2 PT0284	

30	18	18.8	12	2 B61497	seed protein ws-17
31	18	18.8	13	2 G56046	urinary tract ston
32	18	18.8	14	2 A41589	25k elastin-bindin
33	18	18.8	14	2 PA0007	lectin B1 - Psopho
34	18	18.8	14	2 I64815	carbonic anhydrase
35	18	18.8	14	2 PH1347	Ig heavy chain DJ
36	18	18.8	14	2 PH1625	Ig H chain V-D-J r
37	18	18.8	15	2 I49407	placental calcium-
38	18	18.8	15	2 PA0005	lectin A1 - Psopho
39	18	18.8	15	2 PA0006	lectin A3 - Psopho
40	18	18.8	15	2 PA0008	lectin B2 - Psopho
41	18	18.8	15	2 PT0222	Ig heavy chain CDR
42	18	18.8	15	2 PH1310	Ig heavy chain DJ
43	18	18.8	16	2 A59046	alpha-conotoxin MI
44	18	18.8	16	2 PH1778	T cell receptor al
45	18	18.8	17	1 A61339	vespulaekin 1 - e
46	18	18.8	17	1 A05168	conotoxin G [val]
47	18	18.8	17	2 B61491	seed protein ws-2
48	18	18.8	17	2 S69164	ferredoxin al - Ja
49	17	17.7	8	2 PC4131	hypothetical prote
50	17	17.7	8	2 A25836	L-serine dehydrata

## ALIGNMENTS

RESULT 1  
A61612  
allatostatin - tobacco hornworm  
C:Species: Manduca sexta (tobacco hornworm)  
C:Date: 21-Jul-1995 #sequence,revision 21-Jul-1995 #text,change 11-Jul-1997  
C:Accession: A61612  
R:Kramer, S.J.; Toschl, A.; Miller, C.A.; Kataoka, H.; Quistad, G.B.; Li, J.P.; Carne  
Proc. Natl. Acad. Sci. U.S.A. 88, 9458-9462, 1991  
A:Title: Identification of an allatostatin from the tobacco hornworm Manduca sexta.  
A:Reference number: A61612, M0ID:92052112  
A:Accession: A61612  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-15 <KRA>  
C:Keywords: neuropeptide; pyrrolidone; carboxylic acid (Gln) #status experimental  
F:1/Modified site: pyrrolidone,carboxylic acid (Gln)

Query Match 29.2%; Score 28; DB 2; Length 15;  
Best Local Similarity 40.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 FEHCNFDVTT 10  
| | | | |  
Db 4 FROCYNPIS 13

RESULT 2  
PC2215  
fibrinogenolytic proteinase A2 (EC 3.4.21.-) - western diamondback rattlesnake (fregm  
N:Alternate names: alpha-fibrinogenase A2  
C:Species: Crotalus atrox (western diamondback rattlesnake)  
C:Date: 20-Feb-1995 #sequence,revision 20-Feb-1995 #text,change 17-Mar-1999  
C:Accession: PC2215  
R:Hung, C.C.; Chlou, S.H.  
Biochem. Biophys. Res. Commun. 201, 1414-1423, 1994  
A:Title: Isolation of multiple isoforms of alpha-fibrinogenase from the western diamo  
viper.  
A:Reference number: PC2214, M0ID:94296418  
A:Accession: PC2215  
A:Molecule type: protein  
A:Residues: 1-15 <HUN>  
C:Keywords: hydrolase; serine proteinase

Query Match 26.0%; Score 25; DB 2; Length 15;  
Best Local Similarity 42.9%; Pred. No. 6.6e+02;

Matches 3: Conservative 2: Mismatches 2: Indels 0: Gaps 0:  
 QY 2 EHCNND 8  
 : | | :  
 Db 5 DECNINE 11

RESULT 3  
 H56046  
 urinary tract stone matrix protein 10, 42k - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 12-Apr-1995 #sequence\_revision 12-Apr-1995 #text\_change 12-Apr-1995  
 C:Accession: H56046  
 R:Binette, J.P.; Binette, M.B.; Gawinowicz, M.A.; Kendrick, N.  
 submitted to the Protein Sequence Database, February 1995  
 A:Description: Isolation, characterization and sequence of stone proteins.  
 A:Reference number: A56046  
 A:Accession: H56046  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-13 <BIN>

Query Match  
 Best Local Similarity 25.0%; Score 24: DB 2: Length 13:  
 Matches 4: Conservative 1: Mismatches 2: Indels 0: Gaps 0:  
 QY 3 HCNFNDV 9  
 | | | :  
 Db 2 HSYFNDL 8

RESULT 4  
 A11497  
 Transaldolase (EC 2.2.1.2) III - yeast (Pichia jadinii) (fragment)  
 C:Species: Pichia jadinii; Candida utilis  
 C>Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 30-Sep-1993  
 C:Accession: A11497  
 R:Tsolas, O.; Sun, S.C.  
 Arch. Biochem. Biophys. 167, 525-533, 1975  
 A:Title: Isolation of a peptide containing a histidyl-cysteinyl sequence from the acti  
 A:Reference number: A11497; MUID:75145197  
 A:Accession: A11497  
 A:Molecule type: protein  
 A:Residues: 1-9 <TSO>  
 C:Keywords: transferase

Query Match  
 Best Local Similarity 24.0%; Score 23: DB 2: Length 9:  
 Matches 3: Conservative 0: Mismatches 0: Indels 0: Gaps 0:  
 QY 3 HCN 5  
 | | | :  
 Db 4 HCN 6

RESULT 5  
 PS0371  
 hypothetical protein (psac region) - Synechococcus sp. (fragment)  
 C:Species: Synechococcus sp.  
 C>Date: 17-Aug-1992 #sequence\_revision 17-Aug-1992 #text\_change 08-Oct-1999  
 C:Accession: PS0371  
 R:Rhiel, E.; Stiewelt, V.L.; Gasparich, G.E.; Bryant, D.A.  
 Gene 112, 123-128, 1992  
 A:Title: The psac genes of Synechococcus sp. PCC7002 and Cyanophora paradoxa: cloning an  
 A:Reference number: J50694; MUID:92201692  
 A:Accession: PS0371  
 A:Molecule type: DNA  
 A:Residues: 1-14 <RHI>  
 A:Cross-References: GB:M86238; NID:g154574; PIDN:AA27351.1; PID:g552030

Query Match  
 Best Local Similarity 24.0%; Score 23: DB 2: Length 14:  
 Matches 5: Conservative 1: Mismatches 1: Indels 0: Gaps 0:  
 QY 8 DVTTRLR 14  
 | | | | :  
 Db 7 DVTGRUQ 13

RESULT 6  
 A61117  
 somatostatin precursor processing enzyme (EC 3.4.21.-) - American gooselish (fragmen  
 C:Species: Iophus americanus (American gooselish)  
 C>Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 13-Sep-1996  
 C:Accession: A61117  
 R:MacKin, R.B.; Noe, B.D.; Spiess, J.  
 Endocrinology 129, 2263-2265, 1991  
 A:Title: Identification of a somatostatin-14-generating propeptide converting enzyme  
 A:Reference number: A61117; MUID:92007528  
 A:Accession: A61117  
 A:Molecule type: protein  
 A:Residues: 1-17 <MAC>  
 A:Experimental source: pancreatic islets  
 C:Superfamily: kexin; subtilisin homology  
 C:Keywords: hydrolase; serine proteinase

Query Match  
 Best Local Similarity 24.0%; Score 23: DB 2: Length 17:  
 Matches 3: Conservative 3: Mismatches 5: Indels 0: Gaps 0:  
 QY 5 NENDVTRLRE 15  
 | | | : :  
 Db 4 NINDIEVMXD 14

RESULT 7  
 A59042  
 alpha-conotoxin Epi - cone shell (Conus episcopatus)  
 C:Species: Conus episcopatus (bishop's cone)  
 C>Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 13-Aug-1999  
 C:Accession: A59042  
 R:Loughnan, M.; Bond, T.; Atkins, A.; Cuevas, J.; Adams, D.J.; Broxton, N.M.; Livett,  
 J. Biol. Chem. 273, 15667-15674, 1998  
 A:Title: Alpha-conotoxin Epi, a novel sulfated peptide from Conus episcopatus that se  
 A:Reference number: A59042; MUID:98286307  
 A:Accession: A59042  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-16 <LOU>

C:Superfamily: alpha-conotoxin  
 C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neu  
 F:1-16/Product: alpha-conotoxin Epi #status experimental <MAT>  
 F:2-8,3-16/Disulfide bonds: #status experimental  
 F:15/Binding site: sulfate (Tyr) (covalent) #status experimental  
 F:16/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match  
 Best Local Similarity 22.9%; Score 22: DB 2: Length 16:  
 Matches 3: Conservative 1: Mismatches 1: Indels 0: Gaps 0:  
 QY 4 CNFND 8  
 | | | :  
 Db 8 CNMNN 12

RESULT 8  
 G49039  
 T-cell receptor beta chain V-D-J-C region (V beta 4, J beta 2.2) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 30-May-1997  
 C:Accession: G49039

R.Rosenberg, W.M.; Moss, P.A.; Bell, J.I.  
 Eur. J. Immunol. 22, 541-549, 1992  
 A:Title: Variation in human T cell receptor V beta and J beta repertoire: analysis using  
 A:Reference number: A49039; MUID:92164737  
 A:Accession: G49039  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 1-16 <ROS>  
 A:Note: sequence extracted from NCBI backbone (NCBIP:90719)  
 C:Keywords: T-cell receptor

Query Match 22.9%; Score 22; DB 2; Length 16;  
 Best Local Similarity 41.7%; Pred. No. 2.2e+03;  
 Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 4 CNFNDVTRLR 15  
 : | | | | |  
 DB 1 CSVEDGTGRGE 12

RESULT 9  
 T46794  
 Hypothetical protein [imported] - Haloarcula marismortui (fragment)  
 C:Species: Haloarcula marismortui  
 C:Date: 17-Mar-2000 #sequence\_revision 17-Mar-2000 #text\_change 21-Jul-2000  
 C:Accession: T46794  
 R:Arndt, E.  
 FEBS Lett. 267, 193-198, 1990  
 A:Title: Nucleotide sequence of four genes encoding ribosomal proteins from the 'S10 and  
 A:Reference number: S10731; MUID:90336772  
 A:Accession: T46794  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-12 <ARN>  
 A:Cross-references: EMBL:X55311; NID:g43610; PIDN:CAA39015.1; PID:g43611

Query Match 21.9%; Score 21; DB 2; Length 12;  
 Best Local Similarity 80.0%; Pred. No. 2.3e+03;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 13 LRENE 17  
 : | | | | |  
 DB 8 LQENE 12

RESULT 10  
 S08575  
 Botulinum neurotoxin type E - Clostridium botulinum (strain Alaska E-43) (fragment)  
 C:Species: Clostridium botulinum  
 A:Variety: strain Alaska E-43  
 C:Date: 19-Mar-1997 #sequence\_revision 21-Nov-1998 #text\_change 21-Nov-1998  
 C:Accession: S08575  
 R:Schmidt, J.J.; Sachyamoorthy, V.; Dasgupta, B.R.  
 Arch. Biochem. Biophys. 238, 544-548, 1985  
 A:Title: Partial amino acid sequences of botulinum neurotoxins types B and E.  
 A:Reference number: S07128; MUID:85197963  
 A:Accession: S08575  
 A:Molecule type: protein  
 A:Residues: 1-13 <SCH>  
 A:Experimental source: strain Alaska E-43  
 F:1-13/Product: botulinum neurotoxin E light chain (fragment) #status predicted <LIG>

Query Match 21.9%; Score 21; DB 2; Length 13;  
 Best Local Similarity 75.0%; Pred. No. 2.5e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 NEND 8  
 : | | | |  
 DB 7 NYND 10

RESULT 11  
 I52226  
 aldehyde dehydrogenase (EC 1.-.-.-) - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999  
 C:Accession: I52226  
 R:Harper, K.; Jones, D.E.  
 Biochem. Biophys. Res. Commun. 152, 940-947, 1988  
 A:Title: Characterization of a functional recombinant rat liver aldehyde dehydrogenas  
 A:Reference number: I52226; MUID:88209084  
 A:Accession: I52226  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-16 <RES>  
 A:Cross-references: GB:M29320; NID:g202853; PIDN:AAA0722.1; PID:g202854  
 C:Keywords: oxidoreductase

Query Match 21.9%; Score 21; DB 2; Length 16;  
 Best Local Similarity 36.4%; Pred. No. 3.1e+03;  
 Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 5 NENDVTRLR 15  
 : | | | | |  
 DB 3 SLDPTVRARE 13

RESULT 12  
 S71919  
 alcohol dehydrogenase (EC 1.1.1.1) - grass carp (fragment)  
 C:Species: Ctenopharyngodon idella (grass carp)  
 C:Date: 14-Apr-1998 #sequence\_revision 24-Apr-1998 #text\_change 07-May-1999  
 C:Accession: S71919  
 R:Tsui, H.T.; Mock, W.Y.; Lau, K.K.; Fong, W.P.  
 Biochim. Biophys. Acta 1296, 41-46, 1996  
 A:Title: Proteolytic activation of grass carp (Ctenopharyngodon idellus) liver alcohol  
 A:Reference number: S71919; MUID:96350418  
 A:Accession: S71919  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-8 <TSU>  
 A:Note: the source is designated Ctenopharyngodon idellus  
 C:Keywords: NAD; oxidoreductase

Query Match 20.8%; Score 20; DB 2; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 2.2e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 7 NDVTR 12  
 : | | | | |  
 DB 1 SDPTTR 6

RESULT 13  
 A12872  
 transaldolase (EC 2.2.1.2) I - yeast (Pichia jadinii) (fragment)  
 C:Species: Pichia jadinii, Candida utilis  
 C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 30-Sep-1993  
 C:Accession: A12872  
 R:Sun, S.C.; Joris, L.; Tsolas, O.  
 Arch. Biochem. Biophys. 178, 69-78, 1977  
 A:Title: Purification and crystallization of transaldolase isozyme I and evidence for  
 A:Reference number: A12872; MUID:77110646  
 A:Accession: A12872  
 A:Molecule type: protein  
 A:Residues: 1-9 <SUN>  
 C:Keywords: transferase

Query Match 20.8%; Score 20; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 2.2e+03;

Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 HCN 5  
 Db 4 HCB 6

## RESULT 14

PT0077  
 proteochondroitin core protein - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 04-Sep-1998 #sequence\_revision 04-Sep-1998 #text\_change 04-Sep-1998  
 C:Accession: PT0077  
 R:Marcum, J.A.; Thompson, M.A.  
 Biochem. Biophys. Res. Commun. 175, 706-712, 1991  
 A:Title: The amino-terminal region of a proteochondroitin core protein, secreted by aort  
 om human bone.  
 A:Reference number: PT0077; MUID:91207372  
 A:Accession: PT0077  
 A:Status: Preliminary  
 A:Molecule type: protein  
 A:Residues: 1-14 <MAR>  
 C:Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan

Query Match 20.8%; Score 20; DB 2; Length 14;  
 Best Local Similarity 50.0%; Pred. No. 4e+03;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 FEHCNNDVT 10  
 Db 3 FEOKGFWDFT 12

## RESULT 15

S29486  
 GTP-binding protein o-rab3 - electric ray (Discopyge ommata) (fragment)  
 C:Species: Discopyge ommata  
 C:Date: 22-Nov-1993 #sequence\_revision 27-Feb-1997 #text\_change 13-Mar-1997  
 C:Accession: S29486  
 R:Volkmann, W.; Pevsner, J.; Elferink, L.A.; Scheller, R.H.  
 FEBS Lett. 317, 53-56, 1993  
 A:Title: Association of three small GTP-binding proteins with cholinergic synaptic vesic  
 A:Reference number: S29486; MUID:93154521  
 A:Accession: S29486  
 A:Status: Preliminary  
 A:Molecule type: protein  
 A:Residues: 1-14 <VOL>

Query Match 20.8%; Score 20; DB 2; Length 14;  
 Best Local Similarity 80.0%; Pred. No. 4e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 NFNDV 9  
 Db 8 NFNAV 12

## RESULT 16

S32677  
 nitrogene cofactor synthesis protein nifs - Anabaena variabilis (fragment)  
 N:Contains: L-cysteine sulfurtransferase (EC 2.8.1.-)  
 C:Species: Anabaena variabilis  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 03-Dec-1999  
 C:Accession: S32677  
 R:Konerjahn, U.; Boehme, H.  
 submitted to the EMBL Data Library, December 1992  
 A:Description: Cloning and expression in E. coli of the Anabaena.  
 A:Reference number: S32675  
 A:Accession: S32677  
 A:Status: Preliminary  
 A:Molecule type: DNA

A:Residues: 1-15 <MON>  
 A:Cross-references: EMBL:X69898; NID:q296503; PIDN:CAA49523.1; PID:q296506  
 C:Superfamily: nitrogen fixation protein nifs  
 C:Keywords: sulfurtransferase

Query Match 20.8%; Score 20; DB 2; Length 15;  
 Best Local Similarity 42.9%; Pred. No. 4.3e+03;  
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 NDVTTRL 13  
 Db 8 NNATKTV 14

## RESULT 17

PA0009  
 seed storage protein beta-chain 7 - Arabidopsis thaliana (fragment)  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 30-Jun-1992 #sequence\_revision 06-Jan-1995 #text\_change 23-Mar-1995  
 C:Accession: PA0009  
 R:Kano, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.  
 submitted to JIPID, July 1994  
 A:Description: Separation and characterization of Arabidopsis proteins by two-dimensi  
 A:Reference number: PA0001  
 A:Accession: PA0009  
 A:Molecule type: protein  
 A:Residues: 1-15 <KAM>  
 A:Experimental source: seed  
 C:Keywords: seed; storage protein

Query Match 20.8%; Score 20; DB 2; Length 15;  
 Best Local Similarity 42.9%; Pred. No. 4.3e+03;  
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 TTRAREN 16  
 Db 8 TAKIHEN 14

## RESULT 18

D22595  
 bombolitin IV - American common bumblebee  
 C:Species: Bombus pennsylvanicus (American common bumblebee)  
 C:Date: 29-Aug-1987 #sequence\_revision 29-Aug-1987 #text\_change 08-Dec-1995  
 C:Accession: D22595  
 R:Argiolas, A.; Pisano, J.J.  
 J. Biol. Chem. 260, 1437-1444, 1985  
 A:Title: Bombolitins, a new class of mast cell degranulating peptides from the venom  
 A:Reference number: A92504; MUID:85105003  
 A:Accession: D22595  
 A:Molecule type: protein  
 A:Residues: 1-17 <ARG>  
 C:Keywords: amidated carboxyl end; hemolysis; venom  
 F:17/Modified site: amidated carboxyl end (Val) #status experimental

Query Match 20.8%; Score 20; DB 2; Length 17;  
 Best Local Similarity 33.3%; Pred. No. 4.8e+03;  
 Matches 3; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 NFNDVTTRL 13  
 Db 2 NIKDILAKL 10

## RESULT 19

PT0272  
 Ig heavy chain CRD3 region (clone 3-103B) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
 C:Accession: PT0272



R.Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
 J. Exp. Med. 173, 395-407, 1991  
 A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and  
 A:Reference number: P10222; MUID:91108337  
 A:Accession: P10272  
 A:Molecule type: DNA  
 A:Residues: 1-9 <YAM>  
 A:Experimental source: B lymphocyte  
 C:Keywords: heterotrimer; immunoglobulin

Query Match 19.8%; Score 19; DB 2; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 2.2e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 NFND 8  
 DB 2 NMND 5

RESULT 20

peptidylglycine monooxygenase (EC 1.14.17.3), rPAM-5 - rat (fragment)  
 N:Alternate names: peptidylglycine alpha-amidating monooxygenase  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 26-May-1995  
 C:Accession: A42266  
 R:Elipper, B.A.; Green, C.B.; Campbell, T.A.; Stoffers, D.A.; Keutmann, H.T.; Mains, R.E.  
 J. Biol. Chem. 267, 4008-4015, 1992  
 A:Title: Alternative splicing and endoproteolytic processing generate tissue-specific fo  
 A:Reference number: A42266; MUID:92156145  
 A:Accession: A42266  
 A:Molecule type: mRNA  
 A:Status: preliminary  
 A:Residues: 1-9 <EIP>  
 A:Experimental source: pituitary  
 A:Note: sequence extracted from NCBI backbone (NCBIN:82733, NCBIPI:82750)  
 C:Keywords: oxidoreductase

Query Match 19.8%; Score 19; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 2.2e+05;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 HCNFND 8  
 DB 2 HICFND 7

RESULT 21

T-cell receptor beta chain V-D-J region - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997  
 C:Accession: PH0895; PH0896  
 R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.  
 J. Exp. Med. 174, 1467-1476, 1991  
 A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergy  
 A:Reference number: PH0895; MUID:92078857  
 A:Accession: PH0895  
 A:Molecule type: mRNA  
 A:Residues: 1-10 <GOL>  
 A:Experimental source: myelin basic protein-immunized T-cell; clones 15 and 16  
 C:Keywords: T-cell receptor

Query Match 19.8%; Score 19; DB 2; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 4.1e+03;  
 Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 4 CNFNDVTTRL 13  
 DB 1 CASDSSERL 10

RESULT 22  
 aminotransferase chimera DY376 - synthetic (fragment)  
 C:Species: Synthetic  
 C:Date: 28-May-1999 #sequence\_revision 28-May-1999 #text\_change 28-May-1999  
 C:Accession: PC2124  
 R: Miyazawa, K.; Kawaguchi, S.; Okamoto, A.; Kato, R.; Ogawa, T.; Kuramitsu, S.  
 J. Biochem. 115, 568-577, 1994  
 A:Title: Construction of aminotransferase chimeras and analysis of their substrate sp  
 A:Reference number: JX0315; MUID:94334304  
 A:Accession: PC2124  
 A:Molecule type: DNA  
 A:Residues: 1-11 <MY>

C:Comment: This is a chimeric enzyme of Escherichia coli aspartate aminotransferase (C:gene: aspC; tyRB  
 C:Comment: The parental enzymes catalyze the reversible amino group transfer reaction  
 C:Keywords: aminotransferase

Query Match 19.8%; Score 19; DB 4; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 4.5e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 RLRE 15  
 DB 2 RLRE 5

RESULT 23

Ig H chain V-D-J region (clone B-less 120) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 02-Jun-1994 #sequence\_revision 02-Jun-1994 #text\_change 17-Mar-1999  
 C:Accession: PH1627  
 R:Levinson, D.A.; Campos-Torres, J.; Leder, P.  
 J. Exp. Med. 178, 317-329, 1993  
 A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less m  
 A:Reference number: PH1580; MUID:93301609  
 A:Accession: PH1627  
 A:Molecule type: DNA  
 A:Residues: 1-14 <LEV>  
 A:Experimental source: bone marrow pre-B lymphocyte  
 C:Keywords: immunoglobulin

Query Match 19.8%; Score 19; DB 2; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 5.7e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 VTTR 12  
 DB 5 VTTR 8

RESULT 24

Ribosomal protein S20 - Mycobacterium bovis (fragment)  
 C:Species: Mycobacterium bovis  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Jan-1995  
 C:Accession: S36889  
 R:Ohara, N.; Kimura, M.; Higashi, Y.; Yamada, T.  
 FEBS Lett. 331, 9-14, 1993  
 A:Title: Isolation and amino acid sequence of the 30S ribosomal protein S19 from Myco  
 A:Reference number: S36887; MUID:94009653  
 A:Accession: S36889  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-15 <OHA>  
 C:Keywords: protein biosynthesis; ribosome

Query Match 19.8%; Score 19; DB 2; Length 15;  
Best Local Similarity 38.5%; Pred. No. 6.2e+03;  
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 5 NENDVTRLRENE 17  
| | | | |  
Db 2 NIKSQOKRNRRTNE 14

RESULT 25  
PT0095  
H+-transporting ATP synthase (EC 3.6.1.34) gamma chain, mitochondrial - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 21-Aug-1998  
C:Accession: PT0095  
R:Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Morimasa, T.; Tsugita, A.  
Submitted to JIPID, July 1998  
A:Description: Proteome analysis of mouse brain.  
A:Reference number: PT0091  
A:Accession: PT0095  
A:Molecule type: protein  
A:Residues: 1-15 <KAW>  
A:Experimental source: brain, striatum  
C:Keywords: hydrolase; mitochondrion

Query Match 19.8%; Score 19; DB 2; Length 15;  
Best Local Similarity 50.0%; Pred. No. 6.2e+03;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 9 VTTRLR 14  
: | | | |  
Db 6 ITXRLK 11

Search completed: April 2, 2002, 09:25:00  
Job time: 114 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:24:41 : Search time 10.02 seconds

(without alignments)  
62.206 Million cell updates/sec

Title: us-09-020-393b-3\_COPY\_42\_58

Perfect score: 96

Sequence: 1 FEHCNFNDVYTRLENE 17

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 836

Minimum DB seq length: 0

Maximum DB seq length: 17

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	29.2	15	1	ALIS_MANSE
2	23	24.0	9	1	TAL3_PICJA
3	23	24.0	17	1	JHRP_PLAVG
4	22	22.9	10	1	UPA9_HUMAN
5	22	22.9	16	1	CCAL_CONEP
6	21	21.9	15	1	MCA2_RHOOP
7	21	21.9	15	1	UC01_MAIZE
8	21	21.9	15	1	UC13_MAIZE
9	20	20.8	9	1	OXYF_SCYCA
10	20	20.8	9	1	TAL1_PICJA
11	20	20.8	10	1	UR47_HUMAN
12	20	20.8	15	1	NIS1_ANASQ
13	20	20.8	17	1	BOL4_MEGPE
14	19	19.8	12	1	UR2_SCYCA
15	19	19.8	14	1	UN46_CLOPA
16	19	19.8	15	1	ITRB_ALBUJ
17	19	19.8	16	1	FIBA_MUSVI
18	18	18.8	8	1	AL11_CYPDO
19	18	18.8	8	1	LMT2_LOCM1
20	18	18.8	12	1	HCTB_MEGCR
21	18	18.8	13	1	ORCK_ORCLT
22	18	18.8	13	1	UP51_UPEIN
23	18	18.8	14	1	LECB_PSOSC
24	18	18.8	15	1	LECI_PSOSC
25	18	18.8	15	1	LEC2_PSOSC
26	18	18.8	15	1	LEC3_PSOSC
27	18	18.8	16	1	CXA2_CONMA
28	18	18.8	17	1	VESP_VESWC
29	17	17.7	9	1	UPA6_HUMAN
30	17	17.7	13	1	LMT4_LOCM1
31	17	17.7	15	1	CXAL_CONGE
32	17	17.7	16	1	AHL_PRUSE
33	17	17.7	16	1	DBH3_RHILE

34	17	17.7	16	1	FIBA_MELME	P14456	meles meles
35	17	17.7	16	1	LE05_BIOGL	P80744	biomphalari
36	17	17.7	17	1	EFQ_THBAO	Q01697	thermus aqu
37	16	16.7	9	1	OXYT_OCTVU	P80027	octopus vul
38	16	16.7	11	1	T2P1_PROVU	P31031	proteus vul
39	16	16.7	15	1	MM01_RAT	P81563	rattus norv
40	16	16.7	16	1	FIBA_EQUAS	P14449	equus asinu
41	16	16.7	16	1	FOR2_MYRGU	P81437	myrmecia gu
42	16	16.7	16	1	RL6_VIBPR	Q56715	vibriopro
43	16	16.7	16	1	UVSX_BPT6	Q06728	bacterioph
44	16	16.7	16	1	VPR_HVIC4	P05953	human immu
45	16	16.7	16	1	VPR_HVIS3	P19555	human immu
46	16	16.7	17	1	ALYS_MYCPH	P81528	mycobacteri
47	16	16.7	17	1	BOL1_MEGPE	P10521	megabombus
48	15	15.6	6	1	ASP2_LACSN	P82655	lactobacilli
49	15	15.6	9	1	CCAP_CARMA	P38556	carcinus ma
50	15	15.6	9	1	UHA2_HUMAN	P40929	homo sapien

## ALIGNMENTS

RESULT 1	ALIS_MANSE	STANDARD;	PRT;	15 AA.
ID	ALIS_MANSE			
AC	P42559;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	ALATOSTATIN (MAS-AS).			
OS	Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;			
OC	Sphingidae; Sphingidae; Sphinginae; Manduca.			
OX	NCBI_TaxID=7130;			
RN	[1]			
RP	SEQUENCE:			
RC	TISUE-Head;			
RX	MEDLINE=92052112; PubMed=1946359;			
RA	Kramer S.J., Toschl A., Miller C.A., Kataoka H., Quistad G.B.,			
RA	Li J.P., Carney R.L., Schooley D.A.;			
RT	"Identification of an allatostatin from the tobacco hornworm Manduca			
RL	sexta".			
RL	Proc. Natl. Acad. Sci. U.S.A. 88:9458-9462(1991).			
CC	-1- FUNCTION: STRONGLY INHIBITS JUVENILE HORMONE BIOSYNTHESIS IN VITRO			
CC	BY THE CORPORA ALLATA FROM FIFTH-STADIUM LARVAE AND ADULT FEMALES.			
CC	-1- SIMILARITY: BELONGS TO THE ALATOSTATIN FAMILY.			
KW	Neuropeptide. 1			
FT	MOD_RES 1			
FT	MOD_RES 15 AA; 1908 MW; 1605877CDRBC838E CRC64;			
SQ	SEQUENCE			
Query Match	29.2%; Score 28; DB 1; Length 15;			
Best Local Similarity	40.0%; Pred. No. 75;			
Matches	4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;			
QY	1 FEHCNFNDVY 10			
Db	4 FRQCYFNPIS 13			
RESULT 2	TAL3_PICJA	STANDARD;	PRT;	9 AA.
ID	TAL3_PICJA			
AC	P17441;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	01-FEB-1994 (Rel. 28, Last annotation update)			
DE	TRANSALDOLASE III (EC 2.2.1.2) (FRAGMENT).			
OS	Pichia jadinii (Yeast) (Candida utilis).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Pichia.			
OX	NCBI_TaxID=4903;			

RN [1]  
RP SEQUENCE.  
RX MEDLINE=75145197; PubMed=1092268;  
RA Tsolas O., Sun S.C.;  
RT "Isolation of a peptide containing a histidinyl-cysteinyl sequence  
from the active center of transaldolase.";  
RL Arch. Biochem. Biophys. 167:525-533(1975).  
CC -1- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF  
METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.  
CC -1- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE  
3-PHOSPHATE = D-ERYTHROSE 4-PHOSPHATE + D-FRUCTOSE 6-PHOSPHATE.  
CC -1- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY.  
DR PIR: A11497; A11497.  
DR InterPro: IPR001585; Transaldolase.  
DR PROSITE: PS00958; TRANSALDOLASE\_2; PARTIAL.  
DR PROSITE: PS01054; TRANSALDOLASE\_1; PARTIAL.  
KW transferase; Pentose shunt.  
FT NON\_TER 1  
FT NON\_TER 9  
SQ SEQUENCE 9 AA; 1033 MW; 325A31A44EB1E058 CRC64;

Query Match 24.0%; Score 23; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 HCN 5  
| | |  
Db 4 HCN 6

RESULT 3  
JHBP\_PLAVG STANDARD; PRT; 17 AA.  
ID JHBP\_PLAVG  
AC P56675;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DE 15-JUL-1999 (Rel. 38, Last annotation update)  
GN JHBP.  
OS Platyrepia virginialis (Ranchman's tiger moth).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Noctuidae; Arctiidae; Platyrepia.  
OX NCBI\_TaxID=30227;  
RN [1]  
RP SEQUENCE.  
RA Prestwich G.D., Atkinson J.K.;  
RT "Rapid purification and N-terminal amino acid sequence of a  
photoaffinity-labeled juvenile hormone binding protein from an arctiid  
moth larva, Platyrepia virginialis.";  
RL Insect Biochem. 20:801-807(1990).  
CC -1- FUNCTION: PREVENTS JUVENILE HORMONE FROM BEING HYDROLYZED BY  
GENERAL ESTERASES BY COMBINING WITH IT SPECIFICALLY.  
CC -1- SUBCELLULAR LOCATION: SECRETED.  
FT NON\_TER 17  
SQ SEQUENCE 17 AA; 1880 MW; B1205F097732DA8A CRC64;

Query Match 24.0%; Score 23; DB 1; Length 17;  
Best Local Similarity 33.3%; Pred. No. 5.9e+02;  
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 FERHCNENDV 9  
| : | : | :  
Db 4 FDPGSTODI 12

RESULT 4  
UPA9\_HUMAN STANDARD; PRT; 10 AA.  
ID UPA9\_HUMAN  
AC P30095;

DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF PLASMA (SPOT 35) (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Plasma;  
RX MEDLINE=93092937; PubMed=1459097.  
RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquall C.,  
RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
RA Hochstrasser D.F.;  
RT "Plasma protein map: an update by microsequencing.";  
RL Electrophoresis 13:707-714(1992).  
CC -1- MISCELLANEOUS: ON THE 2D-GEL, THE DETERMINED PI OF THIS UNKNOWN  
PROTEIN IS: 7.2. ITS MW IS: 15 KDA.  
DR SWISS-2DPAGE: P30095; HUMAN.  
FT NON\_TER 1  
FT NON\_TER 10  
SQ SEQUENCE 10 AA; 1233 MW; 37AD72B409C681B7 CRC64;

Query Match 22.9%; Score 22; DB 1; Length 10;  
Best Local Similarity 57.1%; Pred. No. 4.9e+02;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 11 TLRENE 17  
| : | : | :  
Db 4 TMFRNE 10

RESULT 5  
CXAL\_CONEP STANDARD; PRT; 16 AA.  
ID CXAL\_CONEP  
AC P56638;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DE 15-JUL-1999 (Rel. 38, Last annotation update)  
GN ALPHA-CONOTOXIN EPI.  
OS Conus episcopatus (Bishop's cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxID=88764;  
RN [1]  
RP X-RAY CRYSTALLOGRAPHY (1.1 ANGSTROMS).  
RX MEDLINE=96376423; PubMed=9708977.  
RA Hu S.H., Loughnan M., Miller R., Weeks C.M., Blessing R.H.,  
RA Alewood P.F., Lewis R.J., Martin J.U.;  
RT "The 1.1-A resolution crystal structure of [Iyr15]EPI, a novel  
alpha-conotoxin from Conus episcopatus, solved by direct methods.";  
RL Biochemistry 37:11425-11433(1998).  
CC -1- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY  
BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACR) AND THUS  
INHIBIT THEM. THIS PEPTIDE BLOCKS MAMMALIAN NICOTINIC  
ACETYLCHOLINE RECEPTORS COMPOSED OF ALPHA-3/BETA-2 AND ALPHA-5  
SUBUNITS.  
CC PDB: 1A0M; 1J-JAN-99.  
DR Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;  
KW Venom; Sulfation; 3d-structure.  
FT DISULFID 2  
FT DISULFID 8  
FT DISULFID 3  
FT MOD\_RES 15  
FT MOD\_RES 16  
FT MOD\_RES 16  
SQ SEQUENCE 16 AA; 1792 MW; C6385F376C9B4C CRC64;

Query Match 22.9%; Score 22; DB 1; Length 16;  
Best Local Similarity 60.0%; Pred. No. 8.1e+02;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 CNFND 8  
1 1  
DB 8 CNMNN 12

## RESULT 6

MCA2\_RHOOP STANDARD; PRT; 15 AA.  
ID MCA2\_RHOOP  
AC P56870;  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DE 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PUTATIVE MALEYLACETATE REDUCTASE II (EC 1.3.1.32) (FRAGMENT).  
OS Rhodococcus opacus (Nocardia opaca).  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacteriaceae; Nocardia; Rhodococcus.  
OX NCBI\_TaxID=37919;  
RN [1]  
RP SEQUENCE.  
RC STRAIN-1CP;  
RX MEDLINE-98324954; PubMed-9657989;  
RA Seibert V., Kourbatova E.M., Golovleva L.A., Schloemann M.;  
RT "Characterization of a maleylacetate reductase encoding region from  
RT Rhodococcus opacus 1CP.";  
RL J. Bacteriol. 180:3503-3508(1998).  
CC -1- CATALYTIC ACTIVITY: 3-OXOADIPATE + NAD(P)(+) = 2-MALEYLACETATE +  
CC NAD(P)H.  
CC -1- PATHWAY: 3-CHLOROCATECHOL DEGRADATION (BETA-KETOADIPATE PATHWAY).  
CC THIS PATHWAY SERVES A VITAL ROLE IN THE BIODEGRADATION OF TOXIC  
CC AROMATIC COMPOUNDS INTRODUCED IN THE ENVIRONMENT BOTH AS NATURAL  
CC PRODUCTS AND AS INDUSTRIAL EFFLUENT.  
CC -1- SIMILARITY: BELONGS TO THE IRON-CONTAINING ALCOHOL DEHYDROGENASE  
CC FAMILY.  
DR InterPro: IPR001670; Fe-ADH.  
DR PROSITE: PS00913; ADH\_IRON\_1; PARTIAL.  
DR PROSITE: PS00060; ADH\_IRON\_2; PARTIAL.  
KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD.  
FT NON\_TER 15  
SQ SEQUENCE 15 AA; 1884 MW; 58DA90DD038F025E CRC64;

Query Match 21.9%; Score 21; DB 1; Length 15;  
Best Local Similarity 80.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FEHCN 5  
1 1 1  
DB 3 FEHCN 7

## RESULT 7

UC01\_MAIZE STANDARD; PRT; 15 AA.  
ID UC01\_MAIZE  
AC P80607;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 113)  
DE (FRAGMENT).  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Coleoptile;  
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,  
RA Perriollet J.-C., Zivy M., de Vienne D.;  
RT "The maize two dimensional gel protein database: towards an integrated  
RT genome analysis program";  
RT Theor. Appl. Genet. 93:997-1005(1996).  
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN

CC PROTEIN IS: 6.3, ITS MW IS: 37.4 KDA.  
DR Maize-2DPAGE; P80607; COLEOPTILE.  
DR MaizeDB: 123920; -;  
FT NON\_TER 1  
FT NON\_TER 15  
SQ SEQUENCE 15 AA; 1844 MW; 6F73A6DDCA0F05A5 CRC64;

Query Match 21.9%; Score 21; DB 1; Length 15;  
Best Local Similarity 44.4%; Pred. No. 1.1e+03;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 FEHCN 5  
1 1 1  
DB 7 YELYNRNDI 15

## RESULT 8

UC13\_MAIZE STANDARD; PRT; 15 AA.  
ID UC13\_MAIZE  
AC P80619;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 243)  
DE (FRAGMENT).  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Coleoptile;  
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,  
RA Perriollet J.-C., Zivy M., de Vienne D.;  
RT "The maize two dimensional gel protein database: towards an integrated  
RT genome analysis program";  
RT Theor. Appl. Genet. 93:997-1005(1996).  
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
CC PROTEIN IS: 6.8, ITS MW IS: 56.9 KDA.  
CC -1- SIMILARITY: TO GLUCOSE-6-PHOSPHATE 1-DEHYDROGENASES.  
DR Maize-2DPAGE; P80619; COLEOPTILE.  
DR MaizeDB: 123946; -;  
FT NON\_TER 1  
FT NON\_TER 15  
SQ SEQUENCE 15 AA; 1739 MW; 02038BE7471AE038 CRC64;

Query Match 21.9%; Score 21; DB 1; Length 15;  
Best Local Similarity 36.4%; Pred. No. 1.1e+03;  
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 7 NDVTRLRENE 17  
1 1 1 1 1  
DB 4 NEFVIRLOXSE 14

## RESULT 9

OXYF\_SCYCA STANDARD; PRT; 9 AA.  
ID OXYF\_SCYCA  
AC P42997;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE PHASVATOCIN.  
OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;  
OC Scyllorhinidae; Scyllorhinus.  
OX NCBI\_TaxID=7830;  
RN [1]  
RP SEQUENCE.

RC TISSUE=pituitary; PubMed=7972045;  
 RX MEDLINE=95062247; Chauvet M.-T., Acher R.;  
 RA "Special evolution of neurohypophyseal hormones in cartilaginous  
 RT fishes: aspartic acid and phasvatoxin, two oxytocin-like peptides  
 RT isolated from the spotted dogfish (*Scyliorhinus caniculus*).";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:11266-11270(1994).  
 CC -1- FUNCTION: DISPLAYS OXYTOCIC ACTIVITY ON RAT UTERUS.  
 CC -1- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.  
 DR InterPro: IPR000981; Neurohypophys\_horm.  
 DR Pfam: PF00220; hormone4; 1  
 DR PROSITE: PS00264; Neurohypophys\_HORM; 1.  
 KW Hormone; Amidation.  
 FT DISULFID  
 FT MOD\_RES 9 6  
 FT SEQUENCE 9 AA; 1016 MW; 17EDD76EB4449DB CRC64;

Query Match 20.8%; Score 20; DB 1; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 1e+05;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 CNEFD 8  
 DB 1 CYFNN 5

RESULT 10  
 ID TALL\_PICJA STANDARD; PRT; 9 AA.  
 AC P17440;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE TRANSALDOLASE I (EC 2.2.1.2) (FRAGMENT).  
 OS Pichia jadinii (Yeast) (Candida utilis).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Pichia.  
 NCBI\_TaxID=4903;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=77110646; PubMed=556924;  
 RA Sun S.C., Joris L., Tsolas O.;  
 RT "Purification of crystallization of transaldolase isozyme I and  
 RT evidence for different genetic origin of isozymes I and III in  
 RT Candida utilis.";  
 RL Arch. Biochem. Biophys. 178:69-78(1977).  
 CC -1- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF  
 CC METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.  
 CC -1- CATALYTIC ACTIVITY: SEDHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE  
 CC -1- 3-PHOSPHATE = D-ERYTHROSE 4-PHOSPHATE + D-FRUCTOSE 6-PHOSPHATE.  
 CC -1- PATHWAY: NONOXIDATIVE PART OF THE PENTOSE-PHOSPHATE PATHWAY.  
 DR PIR: A12872; BELONGS TO THE TRANSALDOLASE FAMILY.  
 DR InterPro: IPR001585; Transaldolase.  
 DR PROSITE: PS00958; TRANSALDOLASE\_2; PARTIAL.  
 DR PROSITE: PS01054; TRANSALDOLASE\_1; PARTIAL.  
 KW Transferrase; Pentose shunt.  
 FT NON\_TER 1 1  
 FT NON\_TER 9 9  
 FT SEQUENCE 9 AA; 1008 MW; 274F31AFOEB1E058 CRC64;

Query Match 20.8%; Score 20; DB 1; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 HCN 5  
 DB 4 HCB 6

RESULT 11

URAT\_HUMAN  
 ID URAT\_HUMAN STANDARD; PRT; 10 AA.  
 AC P34990;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF RED BLOOD CELLS (SPOT 2D-006H0)  
 OS (FRAGMENT).  
 CC Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Erythrocyte;  
 RA Golaz O., Hughes G.J., Frutiger S., Paquet N., Bairoch A.,  
 RA Pasquali C., Sanchez J.-C., Tissot J.-D., Appel R.D., Walzer C.,  
 RA Balant L., Hochstrasser D.F.;  
 CC Submitted (FEB-1994) to the SWISS-PROT data bank.  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 4.77, ITS MW IS: 26 KDA.  
 DR SWISS-2DPAGE; P34990; HDMAN.  
 FT NON\_TER 10 10  
 FT SEQUENCE 10 AA; 1161 MW; F72D82E9C44B1871 CRC64;

Query Match 20.8%; Score 20; DB 1; Length 10;  
 Best Local Similarity 60.0%; Pred. No. 1.1e+03;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 NFNDV 9  
 DB 6 NFNDI 10

RESULT 12  
 ID NISL\_ANASO STANDARD; PRT; 15 AA.  
 AC Q44507;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE CYSTEINE DESULFURASE 1 (EC 4.4.1.-) (NITROGENASE METALLOCLUSTERS  
 DE NIFS1 OR NIFS.  
 OS Anabaena sp. (strain PCC 7937 / ATCC 29413).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.  
 NCBI\_TaxID=1172;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Monnerjahn U., Boelme H.;  
 RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP EXPRESSION PATTERN.  
 RX MEDLINE=96016168; PubMed=7568132;  
 RA Thiel T., Lyons E.M., Erker J.C., Ernst A.;  
 RT "A second nitrogenase in vegetative cells of a heterocyst-forming  
 RT cyanobacterium";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:9358-9362(1995).  
 CC -1- FUNCTION: CATALYZES THE REMOVAL OF ELEMENTAL SULFUR ATOMS FROM  
 CC CYSTEINE TO PRODUCE ALANINE. SEEMS TO PARTICIPATE IN THE  
 CC BIOSYNTHESIS OF THE NITROGENASE METALLOCLUSTERS BY PROVIDING THE  
 CC INORGANIC SULFUR REQUIRED FOR THE FE-S CORE FORMATION (BY  
 CC SIMILARITY).  
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- MISCELLANEOUS: BELONGS TO THE NIFI GENE CLUSTER WHICH IS EXPRESSED  
 CC IN HETEROCYST UNDER ANAEROBIC AND AEROBIC CONDITIONS.  
 CC -1- SIMILARITY: BELONGS TO CLASS-V OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
 CC AMINOTRANSFERASES. NIFS/ISCS SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -



CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eufrosids I; Fabales; Fabaceae; Mimosoideae; Ingeae; Albizia;  
 OX NCBI\_TaxID=3813;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE-Seed;  
 RX MEDLINE=80115605; PubMed=528539;  
 RA Odani S., Ono T., Ikenaka T.;  
 RT "Proteinase inhibitors from a mimosoideae legume, Albizzia  
 julibrissin. Homologues of soybean trypsin inhibitor (Kunitz).";  
 RL J. Biochem. 86:1795-1805(1979).  
 CC -1- FUNCTION: INHIBITS TRYPSIN AND ALPHA-CHYMOTRYPSIN.  
 CC -1- SUBUNIT: HETERODIMER OF AN "A" AND A "B" CHAIN LINKED BY A  
 CC DISULFIDE BOND.  
 CC -1- SIMILARITY: TO BETA CHAINS OF SOME OTHER LEGUMINOUS KUNITZ-TYPE  
 CC INHIBITORS.  
 KW Serine protease inhibitor.  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1705 MW; 53165F7E9C45B4D0 CRC64;

Query Match  
 Best Local Similarity 19.8%; Score 19; DB 1; Length 15;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 EHC 4  
 DB 3 DHC 5

RESULT 17  
 FIBA\_MUSVI  
 ID FIBA\_MUSVI STANDARD; PRT; 16 AA.  
 AC P14458;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DE FIBRINOPEPTIDE A.  
 OS Mustela vison (American mink).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Mustelidae; Mustela.  
 OX NCBI\_TaxID=9667;  
 RN [1]  
 RP SEQUENCE.  
 RA Blomback B., Blomback M., Grondahl N.J.;  
 RT "Studies on fibrinopeptides from mammals.";  
 RL Acta Chem. Scand. 19:1789-1791(1965).  
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION. YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.  
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA, & GAMMA). LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 KW Blood coagulation; Plasma.  
 FT NON\_TER 16 16  
 SQ SEQUENCE 16 AA; 1679 MW; 09432A57919ECF66 CRC64;

Query Match  
 Best Local Similarity 19.8%; Score 19; DB 1; Length 16;  
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 11 TRLENE 17  
 DB 1 TNVASE 7

RESULT 18  
 ALL1\_CYPDPO  
 ID ALL1\_CYPDPO STANDARD; PRT; 8 AA.

AC P82152;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CYDISTATIN 1.  
 OS Cydia pomonella (Coddling moth).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
 OC Tortricidae; Tortricidae; Olethreutinae; Cydia.  
 OX NCBI\_TaxID=82600;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE-Larva;  
 RX MEDLINE=98054539; PubMed=9392829;  
 RA Davey H., Johnson A.H., Maestri J.-L., Scott A.G., Winstanley D.,  
 RA Davey M., East P.D., Thorpe A.;  
 RT "Lepidopteran peptides of the allatostatin superfamily.";  
 RL Peptides 18:1301-1309(1997).  
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 8 8  
 SQ SEQUENCE 8 AA; 934 MW; C82879C45B51F775 CRC64;

Query Match  
 Best Local Similarity 18.8%; Score 18; DB 1; Length 8;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 HCNF 6  
 DB 3 HYNF 6

RESULT 19  
 LMT2\_LOCM1  
 ID LMT2\_LOCM1 STANDARD; PRT; 8 AA.  
 AC P22396;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DE LOCUSTAMYTROPIN 2 (LOM-WT-2).  
 OS Locusta migratoria (Migratory locust).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;  
 OX Acridomorpha; Acridoidea; Acrididae; Locusta.  
 OX NCBI\_TaxID=7004;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Corpora cardiaca;  
 RA Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;  
 RT "Isolation, identification and synthesis of locustamytropin II, an  
 RT additional neuropeptide of Locusta migratoria. Member of the  
 RT cephalomyotropic peptide family.";  
 RL Insect Biochem. 20:479-484(1990).  
 CC -1- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY  
 CC (MYOTROPIC ACTIVITY).  
 CC -1- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.  
 CC InterPro: IPR001484; Pyrokinin.  
 DR PROSITE: PS00539; PYROKININ; 1.  
 KW Neuropeptide; Amidation; Pyrokinin.  
 FT MOD\_RES 8 8  
 SQ SEQUENCE 8 AA; 934 MW; 26341771A9CAAB7B CRC64;

Query Match  
 Best Local Similarity 18.8%; Score 18; DB 1; Length 8;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 8 DVTRL 13  
 DB 3 DTPRL 8



```

RESULT 20
HCYB_MEGCR STANDARD; PRT; 12 AA.
AC 010584:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE HEMOCYANIN B CHAIN (KLH-B) (FRAGMENT).
OS Megathura crenulata (Giant keyhole limpet).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Archaeogastropoda;
OC Patellogastropoda; Megathura.
OX NCB1_TaxID=53429;
RN [1]
RP SEQUENCE.
RX MEDLINE=96208935; PubMed=8829804;
RA Swerdlow R.D., Ebert R.F., Lee P., Bonaventura C., Miller K.I.;
RT "Keyhole limpet hemocyanin: structural and functional
characterization of two different subunits and multimers.";
RL Comp. Biochem. Physiol. 113B:537-548(1996).
CC -1- FUNCTION: HEMOCYANINS ARE COPPER-CONTAINING OXYGEN CARRIERS
OCCURRING FREELY DISSOLVED IN THE HEMOLYMPH OF MANY MOLLUSKS AND
ARTHROPODS.
CC -1- SUBUNIT: DIDECAMERS AND EXTENDED MULTIMERS.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: HEMOLYMPH.
CC -1- BIOTECHNOLOGY: POTENT IMMUNOGEN USED CLASSICALLY AS A CARRIER
PROTEIN FOR HAPTENS AND MORE RECENTLY IN HUMAN VACCINES AND FOR
IMMUNOTHERAPY OF BLADDER CANCER.
CC -1- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY. HEMOCYANIN
CC SUBFAMILY.
DR InterPro: IPR000896; Hemocyanin.
DR PROSITE: PS00209; HEMOCYANIN_1; PARTIAL.
DR PROSITE: PS00210; HEMOCYANIN_2; PARTIAL.
KW Respiratory protein; Oxygen transport; Copper; Glycoprotein;
KM Hemolymph.
FT NON_TER 12 12
SQ SEQUENCE 12 AA; 1345 MW; CBFBEAA44A432412 CRC64;

Query Match 18.8%; Score 18; DB 1; Length 12;
Best Local Similarity 50.0%; Pred. No. 2.8e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 VTTRLEN 16
Db 1 VDTVVRKN 8

RESULT 21
ORCK_ORCLI STANDARD; PRT; 13 AA.
ID ORCK_ORCLI
AC P37086:
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE ORCKOKININ.
OS Orconectes limosus (Spinycheek crayfish).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC Astacidea; Cambaridae; Orconectes.
OX NCB1_TaxID=28379;
RN [1]
RP SEQUENCE.
RX TISSUE=Abdominal nerve cord;
RX MEDLINE=93126144; PubMed=1480511;
RA Stangler J., Hilbich C., Burdzik S., Keller R.;
RT "Orckokin: a novel myotropic peptide from the nervous system of the
crayfish, Orconectes limosus.";
RL Peptides 13:859-864(1992).
CC -1- FUNCTION: MYOTROPIC PEPTIDE, ENHANCES BOTH THE FREQUENCY AND
AMPLITUDE OF SPONTANEOUS HINDGUT CONTRACTIONS.
CC -1- TISSUE SPECIFICITY: ABDOMINAL NERVE CORD AND HINDGUT.
KW Neuropeptide.

```

```

SQ SEQUENCE 13 AA; 1518 MW; 8A31807BA93A40A CRC64;

Query Match 18.8%; Score 18; DB 1; Length 13;
Best Local Similarity 40.0%; Pred. No. 3e+03;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 NFNDV 9
Db 1 NFDEI 5

RESULT 22
UP51_UPEIN STANDARD; PRT; 13 AA.
ID UP51_UPEIN
AC P82036:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE UPERIN 5.1.
OS Uperoleia inundata (Floodplain toadlet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Myobatrachidae;
OC Uperoleia.
OX NCB1_TaxID=104953;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Skin secretion;
RA Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,
RA Adams G.W., Severini C.;
RT "Novel uperin peptides from the dorsal glands of the Australian
floodplain toadlet Uperoleia inundata.";
RL Aust. J. Chem. 49:475-484(1996).
CC -1- FUNCTION: UNKNOWN.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MW=1426; METHOD=FAE.
KW Amphibian skin.
SQ SEQUENCE 13 AA; 1457 MW; 290CA148C384B5A7 CRC64;

Query Match 18.8%; Score 18; DB 1; Length 13;
Best Local Similarity 33.3%; Pred. No. 3e+03;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 FEHCNFNDV 9
Db 1 FQFVNPDSI 9

RESULT 23
LECB_PSOSC STANDARD; PRT; 14 AA.
ID LECB_PSOSC
AC P22584:
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE BASIC LECTIN B1 (FRAGMENT).
OS Psophocarpus scandens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Euphorbia; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
OC Psophocarpus.
OX NCB1_TaxID=3690;
RN [1]
RP SEQUENCE.
RX TISSUE=Seed;
RA Kortt A.;
RT "Isolation and characterization of the lectins from the seeds of
Psophocarpus scandens.";
RL Phytochemistry 27:2847-2855(1988).
CC -1- SUBUNIT: DIMER OF TWO IDENTICAL OR NEARLY IDENTICAL SUBUNITS OF
ABOUT 32000 APPARENT MW.
CC -1- PTM: CONTAINS ABOUT 9% CARBOHYDRATE. HETEROGENEITY IN CARBOHYDRATE

```

CC COMPOSITION MAY RESULT IN DIFFERENT ISOLECTIN FORMS.  
 CC -1- SIMILARITY: WITH P.TETRAGONOLOBUS BASIC LECTINS IN N-TERMINAL  
 CC SEQUENCE, AA COMPOSITION, ERYTHROCYTE AND SUGAR SPECIFICITIES.  
 DR PIR: PA0007; PA0007.  
 DR HSSP: 024313; 1MBL.  
 KW Lectin; Glycoprotein.  
 FT NON\_TER  
 SQ SEQUENCE 14 AA; 1732 MW; D80ACE43B487C549 CRC64;

Query Match  
 Best Local Similarity 18.8%; Score 18; DB 1; Length 14;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 5 NFN 7  
 Db 6 NFN 8

RESULT 24  
 LEC1\_PSO SC STANDARD; PRT; 15 AA.  
 ID LEC1\_PSO SC  
 AC P22582;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE ACIDIC LECTIN A1 (FRAGMENT).  
 OS Psophocarpus scandens.  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC euroids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 CC Psophocarpus.  
 OX NCBI\_TaxID=3890;  
 RN [1]  
 RP TISSUE-Seed;  
 RC "Isolation and characterization of the lectins from the seeds of  
 RL Psophocarpus scandens."  
 RL Phytochemistry 27:2847-2855(1988).  
 CC -1- SUBUNIT: DIMER. THE SUBUNITS SHOW APPARENT MW HETEROGENEITY  
 CC (32000-35000 MW), WHICH MAY RESULT FROM DIFFERENT CARBOHYDRATE  
 CC CONTENT, AA SEQUENCE, OR POLYPEPTIDE LENGTH.  
 CC -1- PTM: CONTAINS 5-6% CARBOHYDRATE.  
 CC -1- SIMILARITY: WITH P.TETRAGONOLOBUS ACIDIC LECTINS IN N-TERMINAL  
 CC SEQUENCE, AA COMPOSITION, ERYTHROCYTE AND SUGAR SPECIFICITIES.  
 DR PIR: PA0005; PA0005.  
 KW Lectin; Glycoprotein.  
 FT NON\_TER  
 SQ SEQUENCE 15 AA; 1743 MW; D6EB9FE9C6B0254D CRC64;

Query Match  
 Best Local Similarity 18.8%; Score 18; DB 1; Length 15;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 5 NFN 7  
 Db 7 NFN 9

RESULT 25  
 LEC2\_PSO SC STANDARD; PRT; 15 AA.  
 ID LEC2\_PSO SC  
 AC P22585;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE BASIC LECTIN B2 (FRAGMENT).  
 OS Psophocarpus scandens.  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC euroids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

CC Psophocarpus.  
 OX NCBI\_TaxID=3890;  
 RN [1]  
 RP TISSUE-Seed;  
 RC "Isolation and characterization of the lectins from the seeds of  
 RL Psophocarpus scandens."  
 RL Phytochemistry 27:2847-2855(1988).  
 CC -1- SUBUNIT: DIMER OF TWO IDENTICAL OR NEARLY IDENTICAL SUBUNITS OF  
 CC ABOUT 32000 APPARENT MW.  
 CC -1- PTM: CONTAINS ABOUT 9% CARBOHYDRATE. HETEROGENEITY IN CARBOHYDRATE  
 CC COMPOSITION MAY RESULT IN DIFFERENT ISOLECTIN FORMS.  
 CC -1- SIMILARITY: WITH P.TETRAGONOLOBUS BASIC LECTINS IN N-TERMINAL  
 CC SEQUENCE, AA COMPOSITION, ERYTHROCYTE AND SUGAR SPECIFICITIES.  
 DR PIR: PA0008; PA0008.  
 KW Lectin; Glycoprotein.  
 FT NON\_TER  
 SQ SEQUENCE 15 AA; 1847 MW; D194CE400C832796 CRC64;

Query Match  
 Best Local Similarity 18.8%; Score 18; DB 1; Length 15;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 5 NFN 7  
 Db 7 NFN 9

Search completed: April 2, 2002, 09:27:44  
 Job time: 183 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:24:11 ; Search time 22.03 Seconds

(without alignments)  
112.875 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58

Perfect score: 96  
Sequence: 1 FEHCNFNDVTRLENE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 3442

Minimum DB seq length: 0  
Maximum DB seq length: 17

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 50 summaries

Database :

1: SPREMBL\_17:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26.0	15	13	Q9PRW3	Q9PRW3 crocalus at
2	25.0	15	13	Q9PRW2	Q9PRW2 crocalus at
3	25.0	9	4	Q14715	Q14715 homo sapien
4	25.0	15	4	P78482	P78482 homo sapien
5	25.0	16	10	Q9AXW2	Q9AXW2 brassica na
6	24.0	16	10	Q9AXV9	Q9AXV9 silene aega
7	23.0	12	10	Q9FSA9	Q9FSA9 silene aega
8	23.0	13	11	P97622	P97622 rattus norv
9	23.0	14	2	055326	055326 synechococ
10	22.9	9	4	085574	085574 homo sapien
11	22.9	16	7	077922	077922 oreochromis
12	22.9	17	12	Q9GEX2	Q9GEX2 human immun
13	21.9	9	5	096417	096417 drosophila
14	21.9	10	6	Q9TQV4	Q9TQV4 equus cabal
15	21.9	10	10	Q9FS93	Q9FS93 silene pent
16	21.9	12	6	Q9TRU1	Q9TRU1 bos taurus
17	21.9	12	10	P82437	P82437 nicotiana t
18	21.9	13	4	Q9UNV6	Q9UNV6 homo sapien
19	21.9	13	10	Q9FSA8	Q9FSA8 silene bac

20	21	21.9	13	10	Q9FS94	Q9FS94 silene pent
21	21	21.9	14	10	Q9FSB1	Q9FSB1 silene aega
22	21	21.9	14	10	Q9FS95	Q9FS95 silene pent
23	21	21.9	14	10	Q9FS91	Q9FS91 silene sedo
24	21	21.9	15	10	Q9FSB0	Q9FSB0 silene aega
25	21	21.9	15	10	Q9FSA5	Q9FSA5 silene cryp
26	21	21.9	15	10	Q9FSA0	Q9FSA0 silene haus
27	21	21.9	15	10	Q9FS99	Q9FS99 silene inte
28	21	21.9	15	10	Q9FS98	Q9FS98 silene laco
29	21	21.9	15	10	Q9FS80	Q9FS80 silene zawa
30	21	21.9	15	10	Q9FE15	Q9FE15 silene sedo
31	21	21.9	15	10	Q9FE14	Q9FE14 silene pent
32	21	21.9	15	10	Q9FE13	Q9FE13 silene aega
33	21	21.9	16	4	Q9BQF2	Q9BQF2 homo sapien
34	21	21.9	16	11	Q63039	Q63039 rattus norv
35	21	21.9	16	12	Q66737	Q66737 equine infe
36	20	20.8	9	2	Q48686	Q48686 lactococcus
37	20	20.8	11	12	Q66874	Q66874 feline call
38	20	20.8	11	12	Q66877	Q66877 feline call
39	20	20.8	13	11	Q54809	Q54809 mus musculu
40	20	20.8	13	12	Q86129	Q86129 vesicular s
41	20	20.8	15	8	Q95952	Q95952 sauromalus
42	20	20.8	15	10	Q9S8P1	Q9S8P1 raphanus sa
43	20	20.8	15	11	Q9QYB7	Q9QYB7 rattus sp.
44	20	20.8	15	11	Q9QY00	Q9QY00 rattus sp.
45	20	20.8	16	5	Q9TWS8	Q9TWS8 lamellibrac
46	19	19.8	8	12	Q84271	Q84271 human papil
47	19	19.8	10	5	Q9TWS1	Q9TWS1 dermatophag
48	19	19.8	12	2	Q9X645	Q9X645 unidentified
49	19	19.8	12	2	Q9X643	Q9X643 klebsiella
50	19	19.8	12	2	Q9X640	Q9X640 citrobacter

#### ALIGNMENTS

RESULT 1  
Q9PRW3 PRELIMINARY; PRT; 15 AA.  
AC Q9PRW3  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE ALPHA-FIBRINOGENASE ISOFORM A2 (FRAGMENT).  
OS Crocalus atrox (Western diamondback rattlesnake).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Chordata; Scleroglossa; Serpentes; Colubroidea;  
OC Viperidae; Crotalinae; Crotalus.  
OX NCBI\_TaxID=8730;  
RN [1]  
RP  
RX MEDLINE=94296418; Pubmed=8024586;  
RA Hung C.C., Chlou S.H.;  
RT "Isolation of multiple isoforms of alpha-fibrinogenase from the  
RT Western diamondback rattlesnake, Crotalus atrox: N-terminal sequence  
RT homology with anurod, an antithrombotic agent from Malaysian viper.";  
RL Biochem. Biophys. Res. Commun. 201:1414-1423(1994).  
SQ SEQUENCE 15 AA; 1640 MW; 03EPE10227CA12DA CRC64;

Query Match 26.0%; Score 25; DB 13; Length 15;  
Best Local Similarity 42.9%; Pred. No. 1.3e+03;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 EHCNFND 8  
Db 5 DECNINE 11

RESULT 2  
Q9PRW2 PRELIMINARY; PRT; 15 AA.  
ID Q9PRW2  
AC Q9PRW2;

DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DE ALPHA-FIBRINOGENASE ISOFORM A3 (FRAGMENT).  
 OS Crotaeus atrox (Western diamondback rattlesnake).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;  
 CC Viperidae; Crotalinae; Crotalus.  
 NC NCBL\_TaxID=8730;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=94296418; PubMed=8024586;  
 RA Hung C.C., Chiu S.H.;  
 RT "Isolation of multiple isoforms of alpha-fibrinogenase from the  
 RT Western diamondback rattlesnake, Crotaeus atrox: N-terminal sequence  
 RT homology with anurod, an antithrombotic agent from Malayan viper.";  
 RL Biochem. Biophys. Res. Commun. 201:1414-1423(1994).  
 SQ SEQUENCE 15 AA; 1656 MW; 03E8E10227D52FDA CRC64;

Query Match  
 Best Local Similarity 26.0%; Score 25; DB 13; Length 15;  
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 EHCNFND 8  
 DB 5 DECININE 11

RESULT 3  
 ID 014715 PRELIMINARY; PRT; 9 AA.  
 AC 014715;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)  
 DE KERATIN 14 (FRAGMENT).  
 GN KRT14.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCBL\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92005680; PubMed=1717157;  
 RA Coulombe P., Hutton M., Letai A., Hebert A., Paller A., Fuchs E.;  
 RT "Point mutations in human keratin 14 genes of epidermolysis bullosa  
 RT simplex patients: genetic and functional analyses.";  
 RL Cell 66:1301-1311(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95072587; PubMed=7526926;  
 RA Yamaniishi K., Matsuki M., Konishi K., Yasuno H.;  
 RT "A novel mutation of Leu122 to Phe at a highly conserved hydrophobic  
 RT residue in the helix initiation motif of keratin 14 in epidermolysis  
 RT bullosa simplex.";  
 RL Hum. Mol. Genet. 3:1171-1172(1994).  
 DR EMBL; D28807; BAA05967.1; -.  
 KW Keratin.  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 9 AA; 1138 MW; BE300AA449C456D6 CRC64;

Query Match  
 Best Local Similarity 25.0%; Score 24; DB 4; Length 9;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 NFND 8  
 DB 4 NFND 7

RESULT 4  
 ID P78482 PRELIMINARY; PRT; 15 AA.  
 AC P78482;  
 DT 01-MAY-1997 (TREMBlrel. 03, Created)  
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE APOLOPROTEIN B48 (FRAGMENT).  
 GN APOB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCBL\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE OF 37-51 FROM N.A.  
 RX MEDLINE=88018019; PubMed=3659919;  
 RA Chen S.H., Habib G., Yang C.Y., Gu Z.W., Lee B.R., Weng S.A.,  
 RA Silberman S.R., Cai S.J., Deslypere J.P., Rosseneu M., Gotto A.M. Jr.,  
 RA Li W.H., Chan L.;  
 RT "Apolipoprotein B-48 is the product of a messenger RNA with an organ-  
 RT specific in-frame stop codon.";  
 RL Science 238:363-366(1987).  
 DR EMBL; M18036; AAA51754.1; -.  
 KW Lipoprotein.  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 15 AA; 1842 MW; 9172790C16559AEB CRC64;

Query Match  
 Best Local Similarity 25.0%; Score 24; DB 4; Length 15;  
 Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 5 NFNDVTRLR 14  
 DB 2 NFNEKLSQLQ 11

RESULT 5  
 ID 09AXW2 PRELIMINARY; PRT; 16 AA.  
 AC 09AXW2;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DE PHYTOCHROME A (FRAGMENT).  
 GN BN-PRYA-1.  
 OS Brassica napus (Rape).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Brassica.  
 NC NCBL\_TaxID=3708;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. STELLAR;  
 RA Foutmann M., Froger N., Brunel D.;  
 RT "Amplified consensus gene markers: Tools designing for a genetic map  
 RT of Arabidopsis-known-function genes in Brassica.";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF229413; AAK00685.1; -.  
 KW Non-terminating transfer RNA.  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 16 AA; 1791 MW; D390F9AD5DDF8F1D CRC64;

Query Match  
 Best Local Similarity 25.0%; Score 24; DB 10; Length 16;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 8 DVTRLRLEN 16  
 DB 5 DIVORMLLEN 13

RESULT 6  
 ID Q9AXV9 PRELIMINARY; PRT; 16 AA.  
 AC Q9AXV9;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE PHYCHROME A (FRAGMENT).  
 GN PHVA-BO-1.  
 OS Brassica oleracea (Cauliflower).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eucosids II; Brassicales; Brassicaceae; Brassica.  
 OX NCBI\_TaxID=3712;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. RAPID CYCLING;  
 RA Foutman M., Froger N., Brunel D.;  
 RT "Amplified consensus gene markers: Tools designing for a genetic map  
 of Arabidopsis-known-function genes in Brassica."  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF229417; AAK00689.1; -.  
 FT NON\_TER 1 16  
 FT NON\_TER 16 16  
 SQ SEQUENCE 16 AA; 1819 MW; 04C0F9AD5DFBF0F CRC64;

Query Match 25.0%; Score 24; DB 10; Length 16;  
 Best Local Similarity 44.4%; Pred. No. 2e+03;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 8 DVTTRLREN 16  
 ID 1 1 11  
 Db 5 DIVORLEN 13  
 RESULT 7  
 ID Q9FSA9 PRELIMINARY; PRT; 12 AA.  
 AC Q9FSA9;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE RNA POLYMERASE II (FRAGMENT).  
 GN RPB2.  
 OS Silene aenea.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllales; Caryophyllaceae; Silene.  
 OX NCBI\_TaxID=49732;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Popp M., Oxelman B.;  
 RT "Inferring the history of the polyploid Silene aenea  
 (Caryophyllaceae) using nuclear and chloroplast DNA sequence data."  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ296130; CAC13014.1; -.  
 FT NON\_TER 1 12  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1304 MW; 83269695B441B2CA CRC64;

Query Match 24.0%; Score 23; DB 10; Length 12;  
 Best Local Similarity 55.6%; Pred. No. 2.2e+03;  
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 6 FNDVTRLR 14  
 ID 1 1 11  
 Db 3 FTDVTSAR 11  
 RESULT 8  
 P97622

ID P97622 PRELIMINARY; PRT; 13 AA.  
 AC P97622;  
 DT 01-MAY-1997 (TREMBlrel. 03, Created)  
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 DE FERRITIN (FRAGMENT).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WISTAR-KYOTO; TISSUE=AORTIC SMOOTH MUSCLE;  
 RA Adams L.A., Werny I., Schwartz S.M.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U75408; AAB19110.1; -.  
 DR HSSP: P02791; IAEW.  
 FT NON\_TER 13 13  
 SQ SEQUENCE 13 AA; 1557 MW; 72A2122AF06D56C4 CRC64;

Query Match 24.0%; Score 23; DB 11; Length 13;  
 Best Local Similarity 37.5%; Pred. No. 2.4e+03;  
 Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 9 VTRLREN 16  
 ID 1 1 11  
 Db 1 MTSQIRON 8  
 RESULT 9  
 ID Q55326 PRELIMINARY; PRT; 14 AA.  
 AC Q55326;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE PUTATIVE ORF1 (FRAGMENT).  
 OS Synechococcus sp.  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.  
 OX NCBI\_TaxID=1131;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=92201692; PubMed=1551590;  
 RA Rhiel E., Strelwalt V.L., Gasparich G.E., Bryant D.A.;  
 RT "The psac genes of Synechococcus sp. PCC7002 and Cynophora paradoxa:  
 cloning and sequence analysis."  
 RL Gene 112:123-128(1992).  
 DR EMBL: M86238; AAA27351.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 1 1  
 SQ SEQUENCE 14 AA; 1590 MW; 9D5226BDCB98BCA CRC64;

Query Match 24.0%; Score 23; DB 2; Length 14;  
 Best Local Similarity 71.4%; Pred. No. 2.3e+03;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 8 DVTTRLR 14  
 ID 1 1 11  
 Db 7 DVTGRLQ 13  
 RESULT 10  
 ID Q95574 PRELIMINARY; PRT; 9 AA.  
 AC Q95574;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE SERUM ALBUMIN (FRAGMENT).  
 GN ALB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FETAL LIVER;
RX MEDLINE=82081882; PubMed=6171778;
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
RT Najarian R.C., Seeburg P.H., Wilson K.L.;
RT "The sequence of human serum albumin cDNA and its expression in E.
RL coli.";
RN Nucleic Acids Res. 9:6103-6114(1981).
RP [2]
RP SEQUENCE FROM N.A.
RC TISSUE=FETAL LIVER;
RX MEDLINE=82105994; PubMed=6275391;
RA Dugalczyk A., Law S.W., Dennison O.E.;
RT "Nucleotide sequence and the encoded amino acids of human serum
RL albumin mRNA.";
RN Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
RP [3]
RP SEQUENCE FROM N.A.
RC TISSUE=FETAL LIVER;
RX MEDLINE=85155492; PubMed=6085063;
RA Urano Y., Sakai M., Watanabe K., Tamaoki T.;
RT "Tandem arrangement of the albumin and alpha-fetoprotein genes in the
RL human genome.";
DR EMBL: L00133; AAC93594.1; -.
FT NON_TER 1
SQ SEQUENCE 9 AA; 1186 MW; 92340B0400440681 CRC64;

```

```

Query Match          22.9%; Score 22; DB 4; Length 9;
Best Local Similarity 50.0%; Pred. No. 4.7e+05;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Oy 10 TTRLRNE 17
Db 2 TMRIRERK 9

```

```

RESULT 11
ID 077922 PRELIMINARY; PRT; 16 AA.
AC 077922;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE MHC CLASS II B LOCUS 16 (FRAGMENT).
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98315113; PubMed=9649539;
RA Mataga-Trillo E., McAndrew B., Vincok V., Zaleska-Rutczynska Z.,
RT Suelthmann H., Figueroa F., Klein J.;
RT "Linkage relationships and haplotype polymorphism among cichlid Mhc
RL class II B loci.";
DR Genetics 149:1527-1547(1998).
DR EMBL: AF050035; AAC41374.1; -.
FT NON_TER 1
SQ SEQUENCE 16 AA; 1893 MW; A56619E46B45F2D CRC64;

```

```

Query Match          22.9%; Score 22; DB 7; Length 16;
Best Local Similarity 75.0%; Pred. No. 4.2e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 4 CNFN 7
Db 10 CDFN 13
RESULT 12
ID 090EX2 PRELIMINARY; PRT; 17 AA.
AC 090EX2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin H.J., Siwak E.B., Hollinger F.B.;
RT "Mutation rate of human immunodeficiency virus type 1 genomic RNA
RL deduced from long term culture of its biological clones.";
DR EMBL: AF178668; AAF04374.1; -.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 17
SQ SEQUENCE 17 AA; 1897 MW; 3346BDC9A99CE117 CRC64;

```

```

Query Match          22.9%; Score 22; DB 12; Length 17;
Best Local Similarity 35.7%; Pred. No. 4.5e+03;
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
Oy 4 CNENDVTTRLRNE 17
Db 2 CTBLNVTITSTEE 15

```

```

RESULT 13
ID 096417 PRELIMINARY; PRT; 9 AA.
AC 096417;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE SXL E1 FORM (FRAGMENT).
GN SXL.
OS Drosophila virilis (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7244;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98337843; PubMed=9671597;
RA Erickson J.W., Cline T.W.;
RT "Key aspects of the primary sex determination mechanism are conserved
RL across the genus Drosophila.";
DR Development 125:3259-3268(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Bell M., Cline T.W.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF046045; AAC97605.1; -.
DR FlyBase: FBgn0016470; Dvir\SL.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1089 MW; 9A8BD1AA9C449CA CRC64;

```

```

Query Match          21.9%; Score 21; DB 5; Length 9;
Best Local Similarity 66.7%; Pred. No. 4.7e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```



OY 5 NENDVT 10  
 11: 11  
 DB 4 NEDVT 9

RESULT 14  
 O9TOV4

ID O9TOV4 PRELIMINARY; PRT: 10 AA.  
 AC O9TOV4;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE POP-VARIANT \* TRANSEPRIN GENE, EXON 17 AND PARTIAL CDS (FRAGMENT).  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Giffard J.M., Brandon R.B., Bell T.K.;  
 RT "Further identification of single nucleotide polymorphisms in the equine transferrin gene.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF185778; AAF05491.1; -;  
 DR EMBL; AF185770; AAF05483.1; -;  
 DR EMBL; AF185771; AAF05484.1; -;  
 DR EMBL; AF185772; AAF05485.1; -;  
 DR EMBL; AF185773; AAF05486.1; -;  
 DR EMBL; AF185774; AAF05487.1; -;  
 DR EMBL; AF185775; AAF05488.1; -;  
 DR EMBL; AF185776; AAF05489.1; -;  
 DR EMBL; AF185777; AAF05490.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 10 AA; 1188 MW; 6DF67DE9D1AEBDB CRC64;

Query Match 21.9%; Score 21; DB 6; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 3.9e+03;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 EHCNENDV 9  
 11: 11  
 DB 3 EACTFHVR 10

RESULT 15  
 O9FS93 PRELIMINARY; PRT: 10 AA.  
 AC O9FS93;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
 DE RNA POLYMERASE II (FRAGMENT).  
 GN RPB2.  
 OS Silene pentellica.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllidae; Caryophyllales; Caryophyllaceae; Silene.  
 OX NCBI\_TaxID=49735;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Popp M., Oxelman B.;  
 RT "Inferring the history of the polyploid Silene aegaea (Caryophyllaceae) using nuclear and chloroplast DNA sequence data.";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ296133; CAC13025.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 10 AA; 1108 MW; CF1AB6D1B2CAB1A9 CRC64;

Query Match 21.9%; Score 21; DB 10; Length 10;

Best Local Similarity 80.0%; Pred. No. 3.9e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 FNDVT 10  
 11: 11  
 DB 3 FNDVT 7

RESULT 16  
 O9TRU1 PRELIMINARY; PRT: 12 AA.  
 AC O9TRU1;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)  
 DE 43 KDA CYANOGEN BROMIDE FRAGMENT PEAK 7.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92132498; PubMed=1734497;  
 RA Veldy O.P., Sletten K., Husby G., Nordstoga K.;  
 RT "Amino acid sequence analyses of non-AA proteins from amyloid fibrils of bovine kidney.";  
 RL Scand. J. Immunol. 35:63-69(1992).  
 SQ SEQUENCE 12 AA; 1257 MW; 0D5C94FDE3B76AAA CRC64;

Query Match 21.9%; Score 21; DB 6; Length 12;  
 Best Local Similarity 80.0%; Pred. No. 4.7e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 FNDVT 10  
 11: 11  
 DB 4 FNDPT 8

RESULT 17  
 P82437 PRELIMINARY; PRT: 12 AA.  
 ID P82437;  
 AC P82437;  
 DT 01-JUN-2000 (TREMblrel. 14, Created)  
 DT 01-JUN-2000 (TREMblrel. 14, Last sequence update)  
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)  
 DE 26 KDA CELL WALL PROTEIN (FRAGMENT).  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN-CV. PETIT HAVANA;  
 RA Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,  
 RA Wojtaszek P., Bolwell G.P.;  
 RT "Proteomic study of secondary cell wall proteins from transformed tobacco culture.";  
 RL Planta 200:0-0(2000).  
 CC -I- SUBCELLULAR LOCATION: CELL WALL.  
 CC -I- TISSUE SPECIFICITY: XYLEM.  
 KM Cell wall.  
 FT NON\_TER 12  
 SQ SEQUENCE 12 AA; 1452 MW; 92B6CDD0F3A45AF1 CRC64;

Query Match 21.9%; Score 21; DB 10; Length 12;  
 Best Local Similarity 80.0%; Pred. No. 4.7e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 7 NDVT 11

Db 7 NDVXT 11

```

RESULT 18
O9UNV6 PRELIMINARY; PRT; 13 AA.
ID O9UNV6:
AC O9UNV6:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE INOSINE MONOPHOSPHATASE 2 (FRAGMENT).
GN IMP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97463449; PubMed=9322233;
RA Yoshikawa T., Turner G., Esterling L.E., Sanders A.R.,
RA Decker-Wadleigh S.D.;
RT "A novel human myo-inositol monophosphatase gene, IMP.18p, maps to a
RT susceptibility region for bipolar disorder.";
RL Mol. Psych. 2:393-397(1997).
DR EMBL: AF085628; AAD22141.1;
DR EMBL: AF085627; AAD22141.1; JOINED.
DR HSP: P29218; IIMB.
FT NON_TER 1
FT NON_TER 13
SQ SEQUENCE 13 AA; 1491 MW; E8154075786DD401 CRC64;

```

Query Match 21.9%; Score 21; DB 4; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 5e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 CNF 6  
 2 CNF 4

```

RESULT 19
O9FSB8 PRELIMINARY; PRT; 13 AA.
ID O9FSB8:
AC O9FSB8:
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE RNA POLYMERASE II (FRAGMENT).
GN RPB2.
OS Silene bacillifera (Berry catchfly) (Cucubalus bacillifer).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllaceae; Silene.
OX NCBI_TaxID=34818;
RN [1]
RP SEQUENCE FROM N.A.
RA Popp M., Oxelman B.;
RT "Inferring the history of the polyploid Silene aegaea
RT (Caryophyllaceae) using nuclear and chloroplast DNA sequence data.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ296139; CAC13015.1;
FT NON_TER 1
FT NON_TER 13
SQ SEQUENCE 13 AA; 1437 MW; D931E2CF1AB6D1B2 CRC64;

```

Query Match 21.9%; Score 21; DB 10; Length 13;  
 Best Local Similarity 80.0%; Pred. No. 5e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 6 FNDVT 10

Db 3 FTDVT 7

```

RESULT 20
O9FS94 PRELIMINARY; PRT; 13 AA.
ID O9FS94:
AC O9FS94:
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE RNA POLYMERASE II (FRAGMENT).
GN RPB2.
OS Silene pentellica.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllaceae; Silene.
OX NCBI_TaxID=49735;
RN [1]
RP SEQUENCE FROM N.A.
RA Popp M., Oxelman B.;
RT "Inferring the history of the polyploid Silene aegaea
RT (Caryophyllaceae) using nuclear and chloroplast DNA sequence data.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ296132; CAC13024.1;
FT NON_TER 1
FT NON_TER 13
SQ SEQUENCE 13 AA; 1437 MW; D931E2CF1AB6D1B2 CRC64;

```

Query Match 21.9%; Score 21; DB 10; Length 13;  
 Best Local Similarity 80.0%; Pred. No. 5e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 6 FNDVT 10  
 3 FTDVT 7

```

RESULT 21
O9FSB1 PRELIMINARY; PRT; 14 AA.
ID O9FSB1:
AC O9FSB1:
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE RNA POLYMERASE II (FRAGMENT).
GN RPB2.
OS Silene aegaea.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllaceae; Silene.
OX NCBI_TaxID=49732;
RN [1]
RP SEQUENCE FROM N.A.
RA Popp M., Oxelman B.;
RT "Inferring the history of the polyploid Silene aegaea
RT (Caryophyllaceae) using nuclear and chloroplast DNA sequence data.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ296122; CAC13007.1;
FT NON_TER 1
FT NON_TER 14
SQ SEQUENCE 14 AA; 1508 MW; 9B0931E2CF1AB6D1 CRC64;

```

Query Match 21.9%; Score 21; DB 10; Length 14;  
 Best Local Similarity 80.0%; Pred. No. 5.4e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 6 FNDVT 10  
 3 FTDVT 7

```

RESULT 22
Q9FS95 PRELIMINARY; PRT; 14 AA.
AC Q9FS95;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE RNA POLYMERASE II (FRAGMENT).
GN RPB2.
OS Silene pentellica.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Caryophyllaceae; Silene.
OX NCBI_TaxID=49735;
RN [1]
RP SEQUENCE FROM N.A.
RA Popp M., Oxelman B.;
RT "Inferring the history of the polyploid Silene aegaea
(RL (Caryophyllaceae) using nuclear and chloroplast DNA sequence data.";
DR EMBL; AJ296131; CAC13023.1; -.
FT NON_TER 1 1
FT SEQUENCE 14 AA; 1508 MW; 9B0931E2CF1AB6D1 CRC64;
SQ

```

```

Query Match
Best Local Similarity 21.9%; Score 21; DB 10; Length 14;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 6 FNDVT 10
DB 3 FTDVT 7

```

```

RESULT 23
Q9FS91 PRELIMINARY; PRT; 14 AA.
AC Q9FS91;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE RNA POLYMERASE II (FRAGMENT).
GN RPB2.
OS Silene sedoides.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Caryophyllaceae; Silene.
OX NCBI_TaxID=39790;
RN [1]
RP SEQUENCE FROM N.A.
RA Popp M., Oxelman B.;
RT "Inferring the history of the polyploid Silene aegaea
(RL (Caryophyllaceae) using nuclear and chloroplast DNA sequence data.";
DR EMBL; AJ296105; CAC13028.1; -.
FT NON_TER 1 1
FT SEQUENCE 14 AA; 1508 MW; 9B0931E2CF1AB6D1 CRC64;
SQ

```

```

Query Match
Best Local Similarity 21.9%; Score 21; DB 10; Length 14;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 6 FNDVT 10
DB 3 FTDVT 7

```

```

RESULT 24
Q9FSB0

```

```

ID Q9FSB0 PRELIMINARY; PRT; 15 AA.
AC Q9FSB0;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE RNA POLYMERASE II (FRAGMENT).
GN RPB2.
OS Silene aegaea.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Caryophyllaceae; Silene.
OX NCBI_TaxID=49732;
RN [1]
RP SEQUENCE FROM N.A.
RA Popp M., Oxelman B.;
RT "Inferring the history of the polyploid Silene aegaea
(RL (Caryophyllaceae) using nuclear and chloroplast DNA sequence data.";
DR EMBL; AJ296124; CAC13009.1; -.
FT NON_TER 1 1
FT SEQUENCE 15 AA; 1616 MW; C96B0931E2CF0AF6 CRC64;
SQ

```

```

Query Match
Best Local Similarity 21.9%; Score 21; DB 10; Length 15;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 6 FNDVT 10
DB 3 FTDVT 7

```

```

RESULT 25
Q9FSAS PRELIMINARY; PRT; 15 AA.
AC Q9FSAS;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE RNA POLYMERASE II (FRAGMENT).
GN RPB2.
OS Silene cryptoneura.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Caryophyllaceae; Silene.
OX NCBI_TaxID=39877;
RN [1]
RP SEQUENCE FROM N.A.
RA Popp M., Oxelman B.;
RT "Inferring the history of the polyploid Silene aegaea
(RL (Caryophyllaceae) using nuclear and chloroplast DNA sequence data.";
DR EMBL; AJ296138; CAC13016.1; -.
FT NON_TER 1 1
FT SEQUENCE 15 AA; 1621 MW; C96B0931E2CF1AB6 CRC64;
SQ

```

```

Query Match
Best Local Similarity 21.9%; Score 21; DB 10; Length 15;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 6 FNDVT 10
DB 3 FTDVT 7

```

```

Search completed: April 2, 2002, 09:27:28
Job time: 197 sec

```

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:22:51 ; Search time 23.72 Seconds  
(without alignments)  
53.088 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58  
Perfect score: 96  
Sequence: 1 FEHCNFNDVTRRLRENE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues  
Total number of hits satisfying chosen parameters: 183508

Minimum DB seq length: 0  
Maximum DB seq length: 17

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database :  
A\_Geneseq\_1101: \*  
1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT: \*  
2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT: \*  
3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT: \*  
4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT: \*  
5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT: \*  
6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT: \*  
7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT: \*  
8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT: \*  
9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT: \*  
10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT: \*  
11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT: \*  
12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT: \*  
13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT: \*  
14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT: \*  
15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT: \*  
16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT: \*  
17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT: \*  
18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT: \*  
19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT: \*  
20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT: \*  
21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT: \*  
22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	39.6	16	22	AAB67300
2	32.5	33.9	16	21	AAV58596
3	32	33.3	15	16	AA83971
4	32	33.3	16	16	AA870779
5	32	33.3	17	18	AA30024
6	32	33.3	17	19	AAW65142
7	30	31.2	8	21	AAV61736
8	30	31.2	9	18	AAW43797
9	30	31.2	9	21	AAV61739
10	30	31.2	12	22	AA897334
11	30	31.2	15	19	AAW41168

12	30	31.2	16	19	AAW41167	Metal-regulated tr
13	29	30.2	9	17	AAW49475	Human leucocyte an
14	29	30.2	14	16	AA871732	Merosin major subu
15	29	30.2	16	21	AA836539	Human serum albumi
16	29	30.2	16	21	AA836546	Human serum albumi
17	28	29.2	6	3	AA820158	Hexapeptide #5 vit
18	28	29.2	9	19	AAI20456	Human microtubule
19	28	29.2	10	22	AAW43078	Myoplasma genital
20	28	29.2	13	21	AAV57800	TRAM-interaction m
21	28	29.2	15	12	AA811465	Leptodera allost
22	27.5	28.6	10	18	AAW32738	Human platelet gly
23	27	28.1	12	20	AA893135	Human growth hormo
24	27	28.1	12	21	AA823282	Hyaluronic acid-bl
25	27	28.1	13	9	AA81812	Sequence of peptid
26	27	28.1	13	9	AA81812	Sequence of peptid
27	27	28.1	13	13	AA822784	HIV-2 env protein
28	27	28.1	14	22	AAE01700	Human gene 3 encod
29	27	28.1	15	22	AA884963	Apoptin-associated
30	27	28.1	16	21	AAV98301	Alpha D peptide de
31	27	28.1	17	21	AAV98399	Alpha D peptide de
32	26.5	27.6	15	17	AAW07941	gp120 peptide HPI0
33	26	27.1	11	21	AA815256	Specific factor VI
34	26	27.1	11	21	AA815259	Specific factor VI
35	26	27.1	12	14	AA835259	D32.39 antibody is
36	26	27.1	12	15	AA85727	Random peptide #7
37	26	27.1	12	17	AA81475	D32.39 monoclonal
38	26	27.1	12	18	AAW5257	Antibody D32.39 ep
39	26	27.1	12	22	AA889315	HIV gp120 protein
40	26	27.1	13	22	AA848698	Mouse IL-3 recepto
41	26	27.1	14	15	AA869361	Stearoyl-VLP(1-14)
42	26	27.1	14	15	AAV94563	Vasodilative intest
43	26	27.1	15	22	AA889312	HIV gp120 protein
44	26	27.1	16	21	AAV79369	Human CXCR3 immuno
45	26	27.1	16	21	AAV56017	Mucor hlemalis end
46	25.5	26.6	12	17	AAV24486	Actinonadura flexu
47	25.5	26.6	12	17	AA894887	Viral integrase in
48	25	26.0	6	18	AAW03262	Cadherin-8 cell ad
49	25	26.0	7	21	AAV61734	Cadherin-8 cell ad
50	25	26.0	10	18	AAW36662	Thrombopoietin rec

## ALIGNMENTS

RESULT 1  
AAB67300 standard; Peptide: 16 AA.  
AC AAB67300;  
DT 20-APR-2001 (first entry)  
XX  
DE Cytokine peptide #6.  
XX  
XX  
KW Prosaposin; cell death; Bcl-XL; brain; cardiac muscle.  
XX  
OS Homo sapiens.  
XX  
XX  
PN EPI072609-A2.  
PD  
PD 31-JAN-2001.  
XX  
PF 30-JUN-2000; 2000EP-0305504.  
PR 30-JUN-1999; 99JP-0185155.  
XX  
XX (SAKA/) SAKANAKA M.  
PA (TANA/) TANAKA J.  
PA (SATO/) SATO K.  
XX  
PI Sakanaka M, Tanaka J, Sato K, Morita F, Sadamoto Y;  
DR WPI; 2001-204263/21.

XX Use of prosaposin-related peptides or derivatives as cytoprotective  
 PT agents, for suppressing apoptosis or apoptosis-like cell death  
 XX  
 PS Disclosure; Page 12; 41pp; English.  
 CC  
 CC The present invention relates to use of a prosaposin-related peptide  
 CC or derivative, in the production of a medicament for use in  
 CC preventing or delaying cell death, or in promoting the expression  
 CC of cell death supporting gene product Bcl-XL. The invention is  
 CC useful for preventing the death of cells e.g. brain cells,  
 CC neurons and cardiac muscle cells, in vitro or ex vivo.  
 XX  
 SQ Sequence 16 AA;

Query Match 39.6%; Score 38; DB 22; Length 16;  
 Best Local Similarity 60.0%; Pred. No. 14;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 EHCNNDVTY 11  
 ||| : : :  
 Db 2 EHCNNDVTY 11

RESULT 2  
 AAY58596  
 ID AAY58596 standard; peptide; 16 AA.  
 XX  
 AC AAY58596;  
 XX  
 DT 10-APR-2000 (first entry)  
 XX  
 DE Toxoplasma gondii antigen SAG1 epitope #8.  
 XX  
 KW SAG1: antigen; toxoplasmosis; subunit vaccine; Pichia pastoris;  
 KM epitope; antibody.  
 OS  
 XX Toxoplasma gondii.  
 PN WO966043-A1.  
 PD 23-DEC-1999.  
 XX  
 PF 08-JUN-1999; 99WO-EP03957.  
 XX  
 PR 12-JUN-1998; 98GB-0012773.  
 PR 15-APR-1999; 99GB-0008564.  
 XX  
 PA (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 PI Biemans R, Bollen A, Haumont M;  
 DR WPI, 2000-106101/09.  
 XX  
 PT Method for production of toxoplasma antigen SAG1 for use in vaccines  
 XX  
 PS Example 1; Page 14; 47pp; English.

XX The invention relates to the recombinant production of the toxoplasma  
 CC antigen SAG1 or a fragment thereof in the yeast Pichia pastoris. SAG1  
 CC is the major surface antigen of Toxoplasma gondii, an obligate  
 CC intracellular protozoan parasite responsible for toxoplasmosis in  
 CC mammals, including humans. SAG1 is anchored to the plasma membrane of  
 CC T. gondii via a GPI (91ycoosylphosphatidylinositol) anchor, which is  
 CC attached to the C-terminal anchor region (residues 308-336). The  
 CC invention provides a truncated SAG1 protein, which lacks the anchor  
 CC region and comprises amino acids 48-307, and a SAG1 protein N-terminally  
 CC fused to a yeast secretion signal peptide. The SAG1 protein and its  
 CC fragments can be used in the manufacture of a subunit vaccine and its  
 CC prevention or treatment of toxoplasmosis in mammals. A live attenuated  
 CC Toxoplasma vaccine is available, but cannot be administered to humans  
 CC due to the risk of reversion of the attenuated strain to a virulent

CC form. Prior art recombinant expression of SAG1 has been attempted in  
 CC Escherichia coli, mammalian cells or S. cerevisiae, but has been  
 CC associated with problems such as misfolding and insolubility. In  
 CC addition, purification of the native protein from tachyzoites is  
 CC difficult and time-consuming, due to the GPI anchor. Sequences  
 CC AAY58595-Y58596 represent peptide epitopes of T. gondii SAG1, which  
 CC were used to raise polyclonal antibodies in an exemplification of the  
 CC present invention.  
 XX  
 SQ Sequence 16 AA;

Query Match 33.9%; Score 32.5; DB 21; Length 16;  
 Best Local Similarity 43.8%; Pred. No. 1e+02;  
 Matches 7; Conservative 2; Mismatches 4; Indels 3; Gaps 1;  
 QY 4 CN---FNDVTTRLREN 16  
 || | : : ||  
 Db 1 CNEKFKDILPKTEN 16

RESULT 3  
 AAR83971  
 ID AAR83971 standard; peptide; 15 AA.  
 XX  
 AC AAR83971;  
 XX  
 DT 21-MAY-1996 (first entry)  
 XX  
 DE Virus infection factor N-terminal peptide from silkworm body fluid.  
 XX  
 KW Amino terminal; silkworm; virus infection factor; promoter;  
 KM recombinant; high yield; body fluid.  
 XX  
 OS Philosamia cynthia ricini.  
 PN JP07252298-A.  
 PD 03-OCT-1995.  
 XX  
 PF 18-NOV-1994; 94JP-0308468.  
 XX  
 PR 22-NOV-1993; 93JP-0314038.  
 XX  
 PA (KATA ) KATAKURA IND CO LTD.  
 XX  
 DR WPI, 1995-371176/48.  
 XX  
 PT Virus infection factor from silkworm body fluid - used to promote  
 PT infection of insect cells with protein-expressing recombinant virus  
 PT in culture medium  
 XX  
 PS Claim 6; Page 2; 11pp; Japanese.  
 XX  
 CC AAR83791 is an amino-terminal peptide of a virus infection factor  
 CC derived from heat-treated silkworm body fluid. The new virus  
 CC infection factor (VIF) is useful for promoting the infection of insect  
 CC cells with recombinant nuclear polyhedrosis virus. The virus having  
 CC been recombined with a useful protein-expressing gene. The insect  
 CC cell can then be cultured and the useful protein expressed at high  
 CC yields. The new VIF has a mol. wt. of approx. 15.2 kDa and an  
 CC isoelectric point of 9 or higher.  
 CC  
 SQ Sequence 15 AA;

Query Match 33.3%; Score 32; DB 16; Length 15;  
 Best Local Similarity 87.5%; Pred. No. 1e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 6 FNDVTTRL 13  
 || |||||

Db 2 fnvvtl1 9

RESULT 4

AAAR70779

ID AAR70779 standard; Peptide; 16 AA.

XX

AC AAR70779;

XX

DT 30-AUG-1995 (first entry)

XX

DE EPO neurotrophic peptide.

XX

XX Erythropoietin; EPO; cytokine; myelination; nervous system;

KW neurotrophic peptide; multiple sclerosis; leukoencephalitis;

KW adrenal leukodystrophy; neuroblastoma.

XX

OS Homo sapiens.

XX

PN WO9503821-A.

XX

PD 09-FEB-1995.

XX

PF 28-JUL-1994; 94MO-US08453.

XX

PR 30-JUL-1993; 93US-0100247.

PR 21-APR-1994; 94US-0232513.

XX

PA (OBRI/) OBRIEN J S.

XX

PI Kishimoto Y, Obrien JS;

XX

DR WPI; 1995-082029/11.

XX

PT Stimulating neural cell out-growth and myelination - with

PT pro:saposin, saposin C or new neurotrophic peptide(s) from

PT cytokine(s), for treating nervous system diseases

XX

PS Disclosure; Page 36; 50pp; English.

XX

XX The peptide given in AAR70773, corresponding to amino acids 8-29 of

CC human saposin-C (AAR70784), promotes neurite outgrowth in vitro.

CC A consensus sequence was determined by comparing the peptide with

CC hematopoietic and neurotrophic cytokines, and neurotrophic peptides

CC (AAR70774-82) were identified in the AB loop of human ciliary

CC neurotrophic factor, interleukins-6, -2, -3 and -gamma,

CC erythropoietin and leukocyte inhibitory factor, and in helix C of human

CC interleukin-1-beta and oncostatin-M. Prosapoin (AAR70783)

CC and saposin-C also promoted nerve cell myelination ex vivo.

XX

XX

SQ Sequence 16 AA;

Query Match 33.3%; Score 32; DB 16; Length 16;

Best Local Similarity 50.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EHCNFDVTT 11

|||:|:|

Db 2 ehcslnenit 11

RESULT 5

AAAM30024

ID AAM30024 standard; peptide; 17 AA.

XX

AC AAM30024;

XX

DT 14-APR-1998 (first entry)

XX

DE Cytokine hEPO AB loop derived peptide.

XX

KW Human; prosapoin; neural disorder; demyelination disorder;

KW neural cell death; inhibition; myelination; neurite outgrowth;

KW neuropathic pain; peripheral nerve disorder; neuroma; mononeuropathy;

KW polynuropathy; dorsal root ganglia; spinal cord; brainstem;

KW thalamus; cytokine.

XX

OS Homo sapiens.

XX

PN WO9732895-A1.

XX

PD 12-SEP-1997.

XX

PF 05-MAR-1997; 97MO-US04143.

XX

PR 05-MAR-1996; 96US-0611307.

XX

PA (REGC ) UNIV CALIFORNIA.

XX

PI O'Brien JS;

XX

DR WPI; 1997-470538/43.

XX

PT Prosapoin-derived peptide - useful for therapy of neural or

PT demyelination disorders in neural tissue

XX

PS Claim 16; Page 56; 69pp; English.

XX

XX The present sequence represents a specifically claimed cytokine derived

CC peptide for use in the following method of the invention. A method

CC has been developed of alleviating or preventing neuropathic pain in a

CC subject, comprising administering an effective amount of an active

CC fragment of prosapoin to the subject. The prosapoin-derived peptide

CC is useful for therapy of neural or demyelination disorders in neural

CC tissue. It can be used to stimulate neurite outgrowth, inhibit neural

CC cell death, promote myelination or inhibit demyelination. The method

CC is used to alleviate neuropathic pain resulting from a peripheral nerve

CC disorder, such as neuroma, nerve compression, crush or stretch and

CC incomplete nerve transection, mononeuropathy or polynuropathy.

CC Alternatively the neuropathic pain results from a disorder of the

CC dorsal root ganglia, spinal cord, brainstem, thalamus or cortex.

XX

XX

SQ Sequence 17 AA;

Query Match 33.3%; Score 32; DB 18; Length 17;

Best Local Similarity 50.0%; Pred. No. 1.3e+02;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EHCNFDVTT 11

|||:|:|

Db 2 ehcslnenit 11

RESULT 6

AAAM6142

ID AAM6142 standard; peptide; 17 AA.

XX

AC AAM6142;

XX

DT 17-NOV-1998 (first entry)

XX

DE Prosapoin receptor agonist #16.

XX

XX Prosapoin; receptor agonist; neuropathic pain; neurite outgrowth;

KW neural cell death; nerve disorder; side effect.

XX

OS Synthetic.

OS Homo sapiens.

XX

PN WO9839357-A1.

XX

PD 11-SEP-1998.

XX

PF 11-SEP-1997; 97MO-US16062.

XX 05-MAR-1997; 97WO-US04143.  
 PR (REGC ) UNIV CALIFORNIA.  
 PA  
 XX  
 PI O'Brien JS;  
 XX  
 DR WPI; 1998-495790/42.  
 XX  
 PT Use of prosaposin receptor agonists - for alleviating neuropathic  
 PT pain, inhibiting sensory or motor neuropathy, or inhibiting neural  
 XX cell death  
 PS  
 XX Claim 13, Page 19; 67pp; English.  
 CC  
 CC The invention relates to prosaposin receptor agonists. Also claimed  
 CC are: (1) methods for alleviating neuropathic pain or inhibiting the  
 CC onset of neuropathic pain, comprising administering a prosaposin receptor  
 CC agonist; (2) inhibiting sensory or motor neuropathy, comprising  
 CC contacting neuronal cells with a composition comprising a prosaposin  
 CC receptor agonist; (3) methods for stimulating neurite outgrowth,  
 CC inhibiting neural cell death, promoting myelination, or inhibiting  
 CC demyelination comprising contacting neuronal cells with a composition  
 CC comprising a prosaposin receptor agonist which has 14-50 amino acids and  
 CC may be used for treatment of a neuropathic pain resulting from peripheral  
 CC nerve disorders (e.g. neuroma, nerve compression, nerve crush, nerve  
 CC stretch, incomplete nerve transection, mononeuropathy or polyneuropathy)  
 CC or disorders of dorsal root ganglia, the spinal cord, the brainstem, the  
 CC thalamus or the cortex. The receptor agonists do not cause undesirable  
 CC side effects. The present sequence represents a prosaposin receptor  
 CC agonist.  
 XX  
 SQ Sequence 17 AA;  
 XX  
 Query Match 33.3%; Score 32; DB 19; Length 17;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EHCNFDVTT 11  
 |||:|:|  
 Db 2 ehcslment 11  
 XX  
 RESULT 7  
 AAY6I736  
 ID AAY6I736 standard; Peptide; 8 AA.  
 XX  
 AC AAY6I736;  
 XX  
 DT 02-MAR-2000 (first entry)  
 XX  
 DE Cadherin-8 cell adhesion recognition cyclic peptide SEQ ID NO:1594.  
 XX  
 KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; Li-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT Disulfide-bond 1..8  
 XX  
 PN WO9957149-A2.  
 XX  
 PD 11-NOV-1999.

XX 05-MAY-1999; 99WO-CA00363.  
 PR  
 XX  
 PR 05-MAY-1998; 98US-0073040.  
 PR 06-NOV-1998; 98US-0187859.  
 PR 20-JAN-1999; 99US-0234395.  
 PR 08-MAR-1999; 99US-0264516.  
 XX  
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX  
 PI Blaschuk OW, Gour BJ, Byers S;  
 XX  
 DR WPI; 2000-038791/03.  
 XX  
 PT New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease -  
 XX  
 PS Claim 42; Page 176; 252pp; English.  
 XX  
 CC The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound  
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue  
 CC in a mammal. They can also be used for treating e.g. psoriasis,  
 CC arthritis, age-related macular degeneration, multiple sclerosis and  
 CC diabetes. The products can also be used for detection and diagnosis and  
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed  
 CC peptides, and AAY64573 to AAY64643 and AAY33183 to AAY33186 represent  
 CC sequences used in the exemplification of the present invention.  
 XX  
 SQ Sequence 8 AA;  
 XX  
 Query Match 31.2%; Score 30; DB 21; Length 8;  
 Best Local Similarity 71.4%; Pred. No. 4.3e+05;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 CNFNDVTT 10  
 |||||  
 Db 1 cqnndvt 7  
 XX  
 RESULT 8  
 AAY43797  
 ID AAY43797 standard; peptide; 9 AA.  
 XX  
 AC AAY43797;  
 XX  
 DT 20-APR-1998 (first entry)  
 XX  
 DE Specific human leukocyte antigen binding peptide #1.  
 XX  
 KW Human leukocyte antigen; HLA; vaccine; bacterial infection; viral;  
 KW cytotoxic T-cell; CTL; immunogenic peptide; cancer.  
 XX  
 OS Synthetic.  
 OS Haemophilus influenzae.  
 XX



PN WO9734617-A1.  
 XX  
 PD 25-SEP-1997.  
 XX  
 PF 21-MAR-1997; 97WO-US04451.  
 XX  
 PR 20-MAR-1997; 97US-0821739.  
 PR 21-MAR-1996; 96US-0013833.  
 XX  
 PA (CYTE-) CYTEL CORP.  
 XX  
 PI Celis E, Grey HM, Kubo RT, Sette A;  
 XX  
 DR WPI; 1997-489250/45.  
 XX  
 PT Specific human leukocyte antigen binding peptide - used in vaccines  
 PT for the treatment and prevention of e.g. bacterial or viral  
 PT infection and cancer  
 PS  
 PS Claim 19; Page 37; 49pp; English.  
 XX  
 CC The present sequence represents a specific example of an immunogenic  
 CC peptide which was used in a new method of inducing a cytotoxic T cell  
 CC (CTL) response against a preselected antigen in a patient. The method  
 CC comprises contacting CTLs from the patient with the immunogenic peptide  
 CC (containing defined motifs) which binds one of the four HLA MHC products  
 CC HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1, the peptide having a dissociation  
 CC constant (Kd) of less than 5 x 10<sup>-7</sup> M. Immunogens are viral, e.g. human  
 CC immunodeficiency virus type I (HIV-1), hepatitis B virus (HBV) and  
 CC hepatitis C virus (HCV) or cancer antigens and are used in vaccines for  
 CC the prevention and treatment of viral infection and cancer. The  
 CC immunogens may be administered to the patient as a nucleic acid encoding  
 CC the peptide (gene vaccine).  
 XX  
 SQ Sequence 9 AA:  
 OY 3 HCNPNDVT 10  
 Db 1 | | | |  
 1 hsnindt 8  
 RESULT 9  
 AAT61739  
 ID AAY61739 standard; Peptide; 9 AA.  
 XX  
 AC AAY61739;  
 XX  
 DT 02-MAR-2000 (first entry)  
 XX  
 DE Cadherin-8 cell adhesion recognition cyclic peptide SEQ ID NO:1597.  
 XX  
 KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW Ob-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; Li-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Disulfide-bond 1..9  
 PN WO957149-A2.

PD 11-NOV-1999.  
 XX  
 PF 05-MAY-1999; 99WO-CA00363.  
 XX  
 PR 05-MAY-1998; 98US-0073040.  
 PR 06-NOV-1998; 98US-0187859.  
 PR 20-JAN-1999; 99US-0234395.  
 PR 08-MAR-1999; 99US-0264516.  
 XX  
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX  
 PI Blaschuk OW, Gour BJ, Byers S;  
 XX  
 DR WPI; 2000-038791/03.  
 XX  
 PT New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease -  
 PS  
 PS Claim 42; Page 176; 252pp; English.  
 XX  
 CC The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MA's can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound  
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue  
 CC in a mammal. They can also be used for treating e.g. psoriasis,  
 CC arthritis, age-related macular degeneration, multiple sclerosis and  
 CC diabetes. The products can also be used for detection and diagnosis and  
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed  
 CC peptides, and AAY64573 to AAY64643 and AA233183 to AA233186 represent  
 CC sequences used in the exemplification of the present invention.  
 XX  
 SQ Sequence 9 AA:  
 OY 4 CNEPDVT 10  
 Db 1 | | | |  
 1 cqindvt 7  
 RESULT 10  
 AAB97334  
 ID AAB97334 standard; peptide; 12 AA.  
 XX  
 AC AAB97334;  
 XX  
 DT 13-AUG-2001 (first entry)  
 XX  
 DE Collagen IV alpha 3 domain epitope peptide #2.  
 XX  
 KW B cell; toxin; antigen specific; antibody mediated disease; virucide;  
 KW immunosuppressive; antiinflammatory; antiallergic; antidiabetic;  
 KW thyromimetic; antithyroid; vasotropic; cardiac; antitumor;  
 KW neuroprotective; antineumatic; antiarthritic; dermatological;  
 KW ophthalmological; nephrotropic; allergy; autoimmune disorder;

KW skin diseases; autoimmune endocrinopathy; vasculitic syndrome;  
 KW cardiovascular disease; immunohematologic disorder; neurologic disease;  
 KW gastrointestinal disease; collagen vascular disease; renal diseases;  
 KW pulmonary disease; infertility disorder; collagen IV;  
 KW Goodpasture syndrome.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200132853-A1.  
 PD 10-MAY-2001.  
 XX  
 PF 12-OCT-2000; 2000WO-US28157.  
 XX  
 PR 29-OCT-1999; 99US-0162464.  
 XX  
 PA (BIOM-) INST APPLIED BIOMEDICINE.  
 XX  
 PI Chaplin JW;  
 XX  
 DR WPI: 2001-316435/33.  
 XX  
 PT B cell clonal toxin useful for treating autoimmune disorders such as  
 PT Grave's disease, myocardial infarction, Crohn's disease, multiple  
 PT sclerosis, comprises a group that causes toxin to be internalized by B  
 cell -  
 XX  
 PS Disclosure: Page 35; 46pp; English.  
 XX  
 CC This invention relates to a B cell clonal toxin. The toxin is made from  
 CC two moieties, the first causes the toxin to be internalized by a B cell,  
 CC and the second is a biologically acceptable toxin. The invention  
 CC includes a method for inactivating/killing an antigen specific B cell. A  
 CC target B cell is contacted with an effective amount of a B cell clonal  
 CC toxin. The method is useful for selective immunosuppression in conditions  
 CC characterized by the presence of an unwanted or deleterious immune  
 CC response. e.g. in the treatment of antigen specific antibody mediated  
 CC disease conditions. Use of the B cell clonal toxin can result in  
 CC immunosuppressive; antiinflammatory; antiallergic; vitruide;  
 CC antidiabetic; thymimetic; antithyroid; vasotropic; cardiant; antiulcer;  
 CC neuroprotective; antinephrotic; antinephrotic; dermatological;  
 CC ophthalmologic; antinephrotic; antinephrotic; antinephrotic;  
 CC useful for treating a host suffering from an antigen specific antibody  
 CC mediated disease condition, where the antigen specific antibody is  
 CC produced by an antigen-reactive B cell population present in a host. The  
 CC toxin is useful for treating allergies, viral disease conditions, and  
 CC autoimmune disorders. Also treated are skin diseases; autoimmune  
 CC endocrinopathies; vasculitic syndromes; cardiovascular disease;  
 CC immunohematologic disorders; gastrointestinal diseases; neurologic  
 CC diseases; collagen vascular diseases; renal diseases; pulmonary diseases;  
 CC and infertility disorders. The present sequence represents a collagen IV  
 CC alpha 3 domain epitope peptide. An antibody response to this antigen is  
 CC implicated in Goodpasture syndrome, a disorder which may be treated  
 CC using the toxin of the invention.  
 XX  
 SQ Sequence 12 AA:

Query Match 31.2%; Score 30; DB 22; Length 12;  
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 CNFNIV 9  
 |||||  
 Db 1 cnvndv 6

RESULT 11  
 AAW41168  
 ID AAW41168 standard; Protein; 15 AA.  
 XX  
 AC AAW41168;  
 XX

DT 20-JUL-1998 (first entry)  
 XX  
 DE Metal-regulated transporter epitope.  
 XX  
 KW MRT; transgenic plant; pollution; metal-deficiency.  
 XX  
 OS Synthetic.  
 OS Arabidopsis thaliana.  
 XX  
 PN WO9745000-A1.  
 PD 04-DEC-1997.  
 XX  
 PF 27-NOV-1996; 96WO-US19065.  
 XX  
 PR 11-OCT-1996; 96CA-2187728.  
 PR 29-MAY-1996; 96US-0018578.  
 XX  
 PA (DART-) DARTMOUTH COLLEGE.  
 PA (MINU) UNIV MINNESOTA.  
 XX  
 PI Elide DJ, Gueriot ML;  
 XX  
 DR WPI: 1998-032230/03.  
 XX  
 PT Novel A. thaliana metal-regulated transporter proteins and related  
 PT DNA - for generating transgenic plants useful for removing  
 PT pollutants from soil or providing nutrient to patients suffering  
 PT from metal-deficiency disorders  
 XX  
 PS Example 16; Page 73; 144pp; English.  
 XX  
 CC The metal regulated transporter epitope spans amino acids 162 through  
 CC 184 of IRT1 (AAV12763). The peptide was used as an antigen to raise  
 CC polyclonal antibodies. These antibodies recognize a protein of  
 CC approximately 33Kda which is only present in iron starved plants.  
 XX  
 SQ Sequence 15 AA:

Query Match 31.2%; Score 30; DB 19; Length 15;  
 Best Local Similarity 45.5%; Pred. No. 2.3e+02;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 7 NDVTRRLRME 17  
 |||||::|::|  
 Db 3 ndvtrlpikedd 13

RESULT 12  
 AAW41167  
 ID AAW41167 standard; Protein; 16 AA.  
 XX  
 AC AAW41167;  
 XX  
 DT 30-JUL-1998 (first entry)  
 XX  
 DE Metal-regulated transporter epitope.  
 XX  
 KW MRT; transgenic plant; pollution; metal-deficiency.  
 XX  
 OS Synthetic.  
 OS Arabidopsis thaliana.  
 XX  
 PN WO9745000-A1.  
 PD 04-DEC-1997.  
 XX  
 PF 27-NOV-1996; 96WO-US19065.  
 XX  
 PR 11-OCT-1996; 96CA-2187728.  
 PR 29-MAY-1996; 96US-0018578.  
 XX

PA (DART-) DARTMOUTH COLLEGE.  
 PA (MIND ) UNIV MINNESOTA.  
 XX  
 PI Elide DJ, Gueriot ML;  
 XX  
 DR WPI: 1998-032230/03.  
 XX  
 PT Novel A. thaliana metal-regulated transporter proteins and related  
 PT DNA - for generating transgenic plants useful for removing  
 PT pollutants from soil or providing nutrient to patients suffering  
 PT from metal-deficiency disorders  
 XX  
 PS Claim 39; Page 105; 144pp; English.  
 XX  
 CC The metal regulated transporter epitope spans amino acids 162 through  
 CC 184 of IRT1 (AAV12763). The peptide was used as an antigen to raise  
 CC polyclonal antibodies. These antibodies recognize a protein of  
 CC approximately 33kDa which is only present in iron starved plants.  
 XX  
 SQ Sequence 16 AA;

Query Match 31.2%; Score 30; DB 19; Length 16;  
 Best Local Similarity 45.5%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 7 NDVTRLRENE 17  
 |||| :|:|:  
 Db 3 ndvtlplkedd 13

RESULT 13  
 AAM49475  
 ID AAM49475 standard; peptide; 9 AA.

AC AAM49475;  
 XX  
 DT 05-JUN-1998 (first entry)  
 XX  
 DE Human leucocyte antigen DQ4 binding peptide #366.  
 XX  
 KW Human leucocyte antigen; HLA-DQ4; combinatorial library;  
 KW autoimmune disease; chronic articular rheumatism.  
 XX  
 OS Synthetic.  
 XX  
 PN JPO8151396-A.  
 PD 11-JUN-1996.  
 XX  
 PF 28-NOV-1994; 94JP-0292657.  
 XX  
 PR 28-NOV-1994; 94JP-0292657.  
 XX  
 PA (TEIJ ) TEIJIN LTD.  
 DR WPI: 1996-329479/33.  
 XX  
 PT HLA-binding oligopeptide and an immuno:regulator contg it - used in  
 PT the treatment of auto:immune disease  
 XX  
 PS Claim 4; Page 43; 61pp; Japanese.  
 XX  
 CC This peptide is an example of a peptide which binds to a human leucocyte  
 CC antigen HLA-DQ4 molecule. The peptide was isolated from a phagemid  
 CC combinatorial library comprising the sequence AAV05953, by screening  
 CC with an HLA-DQ4 molecule. The peptide is used for the treatment of  
 CC autoimmune disease, or especially for treatment of viral diseases.  
 XX  
 SQ Sequence 9 AA;

Query Match 30.2%; Score 29; DB 17; Length 9;

Best Local Similarity 80.0%; Pred. No. 4.3e+05;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 3 HCNFN 7  
 ||||  
 Db 2 hcnfn 6

RESULT 14  
 AAR71732  
 ID AAR71732 standard; Peptide; 14 AA.

AC AAR71732;  
 XX  
 DT 01-MAY-1996 (first entry)  
 XX

DE Merosin major subunit derived peptide #2.

KW Human; 380-400 kD; merosin; major subunit; placenta; striated muscle;  
 KW peripheral nerve; trophoblast; Schwann cell neoplasm; 65 kD subunit;  
 KW 80 kD subunit; merosin polypeptide; merosin subunit; M chain;  
 KW laminin M chain; antigen; antibody; detection; tumour; malignancy;  
 KW neurite outgrowth; inhibitor; cell attachment.

OS Homo sapiens.  
 XX  
 PN MO9508628-A2.  
 PD 30-MAR-1995.  
 XX  
 PF 21-SEP-1994; 94MO-US10730.  
 XX  
 PR 22-SEP-1993; 93US-0125077.  
 XX

PA (LJOL-) LA JOLLA CANCER RES FOUND.

PI Engvall E, Leiyo I;  
 XX

DR WPI: 1995-139597/18.

PT New merosin fragments, corresp. DNA and antibodies - for diagnosing  
 PT tumour malignancy, promoting or inhibiting neurite growth and  
 PT promoting cell attachment.  
 XX

PS Claim 6; Page 45; 65pp; English.

CC The sequences given in AAR71731-32 represent peptides derived from  
 CC the human 380-400 kD merosin major subunit protein. Merosin is an  
 CC isoform of laminin and shows structural and sequence similarity to  
 CC the human laminin A chain. Mature human merosin is 30 amino acids  
 CC larger than the human laminin A chain. Similarly to all laminin chains,  
 CC the merosin protein has distinct domains which are predicted to have  
 CC globular regions, cysteine-rich rod-like regions and helical structures.  
 CC The merosin protein has an apparent mol. wt. of about 800 kD and is  
 CC composed of four polypeptides with molecular weights of 300, 200, 200  
 CC and 80 kD. The 300 kD polypeptide is joined to the 200 kD polypeptides  
 CC by disulphide bonds and the 300 and 80 kD polypeptides comprise the  
 CC major subunit protein given in AAR71729. Merosin is found in placenta,  
 CC striated muscle, peripheral nerve, trophoblasts and human Schwann cell  
 CC neoplasms, among other tissues. The 380-400 major merosin subunit also  
 CC yields a 65 kD subunit. The 380-400 merosin subunit has been designated  
 CC merosin polypeptide, merosin subunit, M chain or laminin M chain. These  
 CC peptides may be used as antigens to raise anti-merosin antibodies.  
 CC These antibodies may be used in the detection of merosin, as the absence  
 CC of merosin in a tumour sample indicates malignancy. Contacting a  
 CC neurone with merosin promotes neurite outgrowth. The merosin polypeptide  
 CC may also be used in contacting inhibitors of neurite outgrowth, thereby  
 CC also promoting the outgrowth. Merosin also promotes cell attachment.  
 CC The merosin gene has been mapped to chromosome 6, more precisely to  
 CC bands 6q22->q23.  
 XX  
 SQ Sequence 14 AA;

Query Match 30.2%; Score 29; DB 16; Length 14;  
Best Local Similarity 38.5%; Pred. No. 3.1e+02;  
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 4 EHCNFNDVTRRLRENE 16  
DB 1 etcfafaeptmrirerk 13

## RESULT 15

AAB36539  
ID AAB36539 standard; Peptide; 16 AA.

AC AAB36539;

DT 07-MAR-2001 (first entry)

DE Human serum albumin (HSA) related (Venezia) peptide sequence.

KW Human serum albumin; HSA.

OS Homo sapiens.

PN CN1266099-A.

PD 13-SEP-2000.

PF 04-MAR-1999; 99CN-0102745.

PR 04-MAR-1999; 99CN-0102745.

PA (MAOI-) MAOJI BIOLOGICAL ENG SCI & TECH CO LTD.

PI Liu Z;

DR WPI; 2000-673206/66.

PT Novel methods for chemical synthesis, expression and recombinant

PT protein production for human serum albumin reformed gene -

PS Disclosure; Fig 2; 85pp; Chinese.

CC The present invention relates to two kinds of DNA sequences of coded  
CC human serum albumin (HSA), i.e. design of structure-modified gene  
CC segment of HSA and artificial total synthesis and a production process  
CC for large-scale production of genetic recombinant HSA by using  
CC methanol, yeast and engineering bacterium, and discovers that the  
CC structure-modified gene can greatly increase the expression quantity  
CC of HSA. The production process can make the structural gene of HSA  
CC obtain high-level expression under the drive of promoter induced by  
CC methanol, and make the HSA expression product secrete into the  
CC fermenting liquor culture medium, and provide reliable test data for  
CC more large-scale pilot-amplification of gene engineering HSA. The  
CC present sequence is given in the exemplification of the present  
CC invention.

SO Sequence 16 AA;

Query Match 30.2%; Score 29; DB 21; Length 16;  
Best Local Similarity 37.5%; Pred. No. 3.6e+02;  
Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 2 EHCNFNDVTRRLRENE 17  
DB 1 etcfafaeptmrirerk 16

## RESULT 16

AAB36546  
ID AAB36546 standard; Peptide; 16 AA.

AC AAB36546;

DT 07-MAR-2001 (first entry)

DE Human serum albumin (HSA) related (Venezia) peptide sequence.

KW Human serum albumin; HSA.

OS Homo sapiens.

PN CN1266100-A.

PD 13-SEP-2000.

PF 04-MAR-1999; 99CN-0102794.

PR 04-MAR-1999; 99CN-0102794.

PA (MAOI-) MAOJI BIOLOGICAL ENG SCI & TECH CO LTD.

PI Liu Z;

DR WPI; 2000-673207/66.

PT Novel methods for the chemical synthesis, expression and recombinant  
PT protein production for human serum albumin reformed gene -

PS Disclosure; Fig 2; 85pp; Chinese.

CC The present invention relates to two kinds of DNA sequences of coded  
CC human serum albumin (HSA), i.e. design of structure-modified gene  
CC segment of HSA and artificial total synthesis and a production process  
CC for large-scale production of genetic recombinant HSA by using  
CC methanol, yeast and engineering bacterium, and discovers that the  
CC structure-modified gene can greatly increase the expression quantity  
CC of HSA. The production process can make the structural gene of HSA  
CC obtain high-level expression under the drive of promoter induced by  
CC methanol, and make the HSA expression product secrete into the  
CC fermenting liquor culture medium, and provide reliable test data for  
CC more large-scale pilot-amplification of gene engineering HSA. The  
CC present sequence is given in the exemplification of the present  
CC invention.

SO Sequence 16 AA;

Query Match 30.2%; Score 29; DB 21; Length 16;  
Best Local Similarity 37.5%; Pred. No. 3.6e+02;  
Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 2 EHCNFNDVTRRLRENE 17  
DB 1 etcfafaeptmrirerk 16

## RESULT 17

AAP20158  
ID AAP20158 standard; Protein; 6 AA.

AC AAP20158;

DT 01-DEC-1992 (first entry)

DE Hexapeptide #5 with learning response improvement properties.

KW asthenia; memory loss; learning; senility; metal complex.

OS Synthetic.

PN

FT Key Location/Qualifiers

FT Misc-difference 5 /note="D-form residue"

FT Disulfide-bond 1..4

[illegible]

XX	(UUT-) RIJKSUNIV UTRECHT.
PA	(ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
PA	(UUTO-) UNIV ROTTERDAM ERASMUS.
PI	Burbach JPH, Grosveld FG, Van Leeuwen FW;
XX	WPI: 1998-609901/51.
DR	N-PSDB; AAX75757.
XX	
PT	Diagnosing disease by detecting frameshift mutations in RNA or
PT	corresponding protein mutations - used to diagnose cancer and
PT	neurological diseases, particularly Alzheimer's disease, and also
PT	for treatment and prevention with specific ribozymes or wild-type
PT	RNA
XX	
PS	Disclosure: Figure 6; 258pp; English.
XX	
CC	This invention describes a novel method for the diagnosis of a disease
CC	caused by, or associated with, an RNA molecule that has a frameshift
CC	mutation. The method is used to diagnose age-related diseases, especially
CC	cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
CC	disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
CC	multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
CC	and many others listed) or susceptibility to these disorders. The method
CC	allows a definitive diagnosis of Alzheimer's disease in living patients,
CC	at an early stage. It is based on the observation that disease may be
CC	caused by mutations in RNA rather than DNA. The invention describes the
CC	use of neuronal system RNA molecules, specifically proteins including
CC	beta-amyloid precursor protein (beta-ApP), the microtubule associated
CC	proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC	associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC	neurofilament-F, presenilin 1, presenilin II, glial fibrillary acidic
CC	protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC	2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
CC	protein-C (HMGP-C) and neuroendocrine specific protein A.
XX	
SO	Sequence 9 AA:
QY	1 FEHCNFN 7
DB	1 fhncsan 7
RESULT 19	
AA043078	
ID	AA043078 standard; Peptide: 10 AA.
XX	AA043078;
AC	
XX	
DT	22-OCT-2001 (first entry)
XX	
DE	Mycoplasma genitalium Intermolecular complementary peptide, SEQ ID 387.
XX	
KW	Mycoplasma genitalium; complementary peptide; ligand;
XX	protein-protein interaction; drug design; Intermolecular;
XX	intramolecular.
OS	Mycoplasma genitalium.
PN	WO200142278-A2.
XX	
PD	14-JUN-2001.
PF	13-DEC-2000; 2000WO-GB04778.
XX	
PR	13-DEC-1999; 99GB-0029466.

XX (PROT-) PROTEOM LTD.  
 PA  
 XX  
 PI Roberts GW, Heal JR;  
 XX  
 DR WPI: 2001-514238/56.  
 XX  
 PT Complementary peptide ligands as reagents and drugs for drug discovery  
 PT programs and as lead ligands to facilitate drug design and development,  
 PT are generated from microbial genome sequences -  
 XX  
 PS Example 2; Page 103; 161pp; English.  
 XX  
 CC The present sequence is one of a large number of complementary peptide  
 CC ligands generated from Mycoplasma genitalium genome sequences. These  
 CC specific complementary peptides interact with their relevant target  
 CC proteins encoded by the microbial genome. They are capable of  
 CC antagonising or agonising specific interaction of a protein with  
 CC another protein or receptor and are thus useful as reagents and drugs,  
 CC and as lead ligands to facilitate drug design and development. They  
 CC are useful as tools for functional genomic studies, reagents for the  
 CC configuration of high-throughput screens, as a starting point for  
 CC medicinal chemistry manipulation, for peptide mimetics and as  
 CC therapeutic agents. The analysis and acquisition of peptide sequences  
 CC facilitates understanding of protein-protein interactions. The method  
 CC allows for analysis of an entire database at a time, thus overcoming  
 CC sampling problems. The set of complementary peptides includes both  
 CC intermolecular (between proteins) and intramolecular (within a  
 CC protein) sequences.  
 XX  
 SQ Sequence 10 AA;  
 XX  
 Query Match 29.2%; Score 28; DB 22; Length 10;  
 Best Local Similarity 55.6%; Pred. No. 3.2e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 6 FNDVTRLR 14  
 Db 1 fsdltlrv 9  
 XX  
 RESULT 20  
 AAY57800  
 ID AAY57800 standard; peptide; 13 AA.  
 XX  
 AC AAY57800;  
 XX  
 DT 20-MAR-2000 (first entry)  
 XX  
 DE TRAM-interaction motif E2F.  
 XX  
 KW Transcriptional adaptor motif; TRAM; TRAM interaction motif; TRIM;  
 KW creb binding protein; transcriptional regulation; cytosolic; antiviral;  
 KW cell cycle inhibitor; viral transcription inhibitor; cancer; tumour;  
 KW viral disease; viral infection; cell cycle; apoptosis; growth arrest.  
 OS Unidentified.  
 XX  
 PN WO9961608-A2.  
 XX  
 PD 02-DEC-1999.  
 XX  
 PF 26-MAY-1999; 99WO-GB01668.  
 XX  
 PR 26-MAY-1998; 98GB-0011303.  
 PR 05-JAN-1999; 99GB-0000157.  
 XX  
 PA (MOLE-) INST MOLECULAR & CELL BIOLOGY.  
 PA  
 PI O'Connor MJ, Zimmermann H;  
 PI  
 DR WPI: 2000-072620/06.

XX Novel polypeptides cells useful for treating viral disease and cancer  
 PT  
 XX  
 PS Example 1; Fig 2; 73pp; English.  
 XX  
 CC The present invention describes a polypeptide comprising a  
 CC transcriptional adaptor motif (TRAM) or a TRAM-interaction motif  
 CC (TRIM) (I). (I) can be used for identifying compounds (II) comprising a  
 CC polypeptide capable of disrupting an interaction between a TRAM sequence  
 CC and/or a TRIM sequence. Polypeptides which bind to a TRAM sequence or a  
 CC TRIM sequence are identified by incubating, the polypeptide with (I) and  
 CC determining if the polypeptide interacts with (I). Compound (II) is  
 CC useful for preparing a pharmaceutical composition and for disrupting an  
 CC interaction between TRAM sequence and TRIM sequence in vitro, thereby  
 CC inhibiting viral transcription or cell cycle progression in mammalian  
 CC cells especially cancer cell. Compounds which disrupt interaction  
 CC between TRAM/TRAM containing polypeptides can be used therapeutically to  
 CC prevent or treat viral diseases and tumours. The polypeptides reduce  
 CC susceptibility of cells to viral infection and regulate cell cycle  
 CC including apoptosis and growth arrest and can be used to produce  
 CC antibodies against the TRIM or TRAM sequences. HPV types associated  
 CC with high risk or low risk of cervical cancer can be distinguished  
 CC based on the ability of B6 polypeptides to bind to creb binding protein  
 CC (CBP) TRAM sequence. The present sequence represents a TRIM from an  
 CC example from the present invention.  
 XX  
 SQ Sequence 13 AA;  
 XX  
 Query Match 29.2%; Score 28; DB 21; Length 13;  
 Best Local Similarity 57.1%; Pred. No. 4.2e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 CNEPDVVT 10  
 Db 3 cdfgdlr 9  
 XX  
 RESULT 21  
 AAR11465  
 ID AAR11465 standard; peptide; 15 AA.  
 XX  
 AC AAR11465;  
 XX  
 DT 11-JUN-1991 (first entry)  
 XX  
 DE Lepidoptera allostatin polypeptide analogue.  
 XX  
 KW LAS; juvenile hormone; insecticide.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1.1  
 FT /label= OTHER  
 FT /note= "pyroglutamine"  
 XX  
 PN EP421935-A.  
 XX  
 PD 10-APR-1991.  
 XX  
 PF 28-SEP-1990; 90EP-0810745.  
 XX  
 PR 02-OCT-1989; 89US-0416097.  
 XX  
 PA (SANO ) SANDOZ LTD.  
 PA  
 PI Kramer SJ, Miller CA, Schooley DA, Toschi AG;  
 PI  
 DR WPI: 1991-104135/15.  
 XX  
 PT Polypeptide exhibiting lepidoptera allostatin like activity - and  
 PT inhibiting juvenile hormone synthesis, used as insecticide and  
 PT insect population control agent

PS Claim 6; page 20; 21pp; English.  
 XX This lepidoptera allostatin (LAS) polypeptide analogue is an  
 CC example of a highly generic polypeptide sequence. It exhibits  
 CC LAS-like activity and inhibits juvenile hormone (JH)-I, -II and  
 CC -III activity in lepidoptera. It is used as an insecticide or  
 CC insect population control agent.  
 XX

Sequence 15 AA;

Query Match 29.2%; Score 28; DB 12; Length 15;  
 Best Local Similarity 40.0%; Pred. No. 4.8e+02;  
 Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 FEHCNFNDVY 10  
 | | | | |  
 Db 4 frqcyInpIs 13

RESULT 22

AAW32738  
 ID AAW32738 standard; peptide; 10 AA.

XX AAW32738;

AC AAW32738;

XX 30-JAN-1998 (first entry)

DE Human platelet glycoprotein Ib/IX complex epitope mimotope.

XX Human platelet glycoprotein Ib/IX complex epitope; probe;

XX monoclonal antibody S2-2; anti mimotope; binding site; antibody;

XX platelet adhesion; aggregation; agglutination; von Willebrand factor;

XX anti thrombotic drug; ristocetin.

XX Synthetic.

OS Homo sapiens.

PN WO9718236-A1.

XX 22-MAY-1997.

XX 08-NOV-1996; 96WO-US17882.

XX 13-NOV-1995; 95US-0556597.

XX (UYNV ) UNIV NEW YORK STATE RES FOUND.

PI Lyle VA, Miller JL;

XX WPI; 1997-289227/26.

XX peptide which mimics human platelet glycoprotein Ib/IX complex

XX epitope - also anti-mimotope molecule capable of binding to the

XX peptide and useful to modulate platelet adhesion, aggregation or

XX agglutination

XX Claim 8; Page 99; 115pp; English.

XX The present sequence represents a mimotope which functionally mimics a

XX binding site for a monoclonal antibody (Mab) which recognises an epitope

XX within the human platelet glycoprotein (gp) Ib/IX complex. More

XX specifically the Mab is S2-2. The mimotope peptide can be used for

XX raising antibodies, as probes to search for anti-mimotopes and to

XX neutralise the inhibitory activity of those antibodies which recognise

XX the epitope which is mimicked by the peptide. Anti-mimotopes can be used

XX to modulate the adhesion, aggregation or agglutination of platelets by

XX affecting von Willebrand factor interaction with the platelets through

XX the gp Ib/IX receptor, e.g. as an anti-thrombotic drug which inhibits

XX the ristocetin induced aggregation of platelets.

XX Sequence 10 AA;

Query Match 28.6%; Score 27.5; DB 18; Length 10;  
 Best Local Similarity 46.2%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 3; Gaps 1;

QY 2 EHCNFNDVYTRLR 14  
 : | : : | | | |  
 Db 1 khcs---ihtrlr 10

RESULT 23

AAW93135  
 ID AAW93135 standard; peptide; 12 AA.

XX AAW93135;

AC AAW93135;

XX 20-MAY-1999 (first entry)

DE Human growth hormone releasing factor peptide #12.

XX Growth hormone releasing factor; GRF; human; treatment; osteoporosis;

XX hypothalamic pituitary dwarfism; burn; renal failure; infection; wound;

XX non-union bone fracture; lactation failure; female infertility; cachexia;

XX cancer; anabolic; catabolic; T-cell immunodeficiency; tumour; marker;

XX neurodegenerative condition; post-surgery.

XX Homo sapiens.

OS US5854216-A.

PN 29-DEC-1998.

XX 24-JUL-1996; 96US-0685357.

XX 24-JUL-1996; 96US-0685357.

XX 23-SEP-1994; 94US-0312244.

XX (UYMO-) UNIV MONTREAL.

PA Gaudreau P;

XX WPI; 1999-094967/08.

XX New polypeptide derivatives are growth hormone releasing factor

XX receptor agonists - useful for treatment of e.g. dwarfism, burns,

XX osteoporosis, renal failure, infections, infertility, cachexia and

XX tumours

XX Disclosure; Column 39-40; 26pp; English.

XX This invention describes novel polypeptides of formula RaxRb where

XX X = C(=O), HNC(=S), HNC(=O)CH2 or CH2; Ra = fluorescein, rhodamine,

XX Texas red, a 'body' (RTM), 'Cascade Blue' (RTM), coumarin,

XX phycoerythrin, eosin or roseamine; and Rb = a polypeptide of 29 or

XX 30 amino acids selected from 26 sequences given in the specification.

XX Such peptides are useful for the treatment of hypothalamic pituitary

XX dwarfism, burns, osteoporosis, renal failure, non-union bone fracture,

XX acute or chronic debilitating illness or infection, wounds, lactation

XX failure, infertility in women, cachexia in cancer patients, anabolic

XX and/or catabolic problems, T-cell immunodeficiencies, neurodegenerative

XX conditions or growth hormone releasing factor (GRF) receptor-dependent

XX tumours and for reduction of the incidence of post-surgical problems

XX and as markers for GRF receptors.

XX Sequence 12 AA;

QY 8 DVYTRLRNE 17  
 | | | | | | | | | |

Db 1 dftltjrdde 10

RESULT 24  
AAB23282  
ID AAB23282 standard; peptide; 12 AA.  
XX  
AC AAB23282;  
XX  
DT 02-FEB-2001 (first entry)  
XX  
DE Hyaluronic acid-binding phage display library peptide, SEQ ID NO:1.  
XX  
KW Hyaluronic acid-binding peptide; glycosaminoglycan; autoimmune disease;  
KW HA-CD4 mediated immune cell migration inhibitor; inflammatory disease;  
KW leukocyte infiltration inhibitor; lupus erythematosus;  
KW rheumatoid arthritis; graft-versus host disease; infectious disease;  
KW hyaluronic acid associated disease; stroke; myocardial infarction;  
KW cancer; brain tumour; oedema; glioblastoma multiforme; HIV;  
KW contraceptive; phage display library peptide.  
XX  
OS Synthetic.  
XX  
EN WO200057896-A1.  
XX  
PD 05-OCT-2000.  
XX  
PF 22-MAR-2000; 2000WO-US07813.  
XX  
PR 26-MAR-1999; 99US-0126475.  
XX  
PA (TEXA ) UNIV TEXAS SYSTEM.  
XX  
PI Takashima A, Mummert ME, Mohammadzadeh M,  
XX WPI; 2000-672524/65.  
XX  
DR Novel composition useful as inhibitor or enhancer of  
XX glycosaminoglycan-mediated processes, for treating or preventing  
XX infections, inflammatory diseases and cancers -  
XX  
PS Claim 4; Page 26; 53pp; English.  
XX  
XX The invention relates to a novel composition comprising an active  
XX ingredient which specifically and selectively binds to a natural  
XX glycosaminoglycan to alter its normal interactions. The invention also  
XX relates to hyaluronic acid (HA)-binding phage display library peptides  
XX (AAB23281-823286) which are able to modulate carbohydrate-mediated  
XX interactions, and a method of identifying peptides able to modulate  
XX carbohydrate-mediated interactions. The composition and peptides  
XX modulate HA-CD4 mediated migration of immune cells and inhibit of  
XX leukocyte infiltration. The composition of the invention is useful for  
XX inhibiting the normal function of natural glycosaminoglycans, preferably  
XX hyaluronic acid, chondroitin sulphates A, B, and C, dermatan sulphate,  
XX heparin, keratan sulphate, keratansulphate, chitin, chitosan 1 or chitosan  
XX 2. The HA-binding peptides are useful for inhibiting an immune reaction,  
XX and may be used in treatment and/or prevention of inflammatory or  
XX autoimmune diseases such as lupus erythematosus, rheumatoid arthritis and  
XX graft-versus host disease. The HA-binding peptides are also useful for  
XX treating diseases associated with hyaluronic acid, such as stroke and  
XX myocardial infarction, and/or as an anti-cancer agents, either alone or  
XX in combination with chemotherapeutic agents. They can be used to treat  
XX oedema associated with brain tumours, particularly oedema associated  
XX with glioblastoma multiforme. The peptides are also useful for treating,  
XX ameliorating and/or preventing infectious diseases such as HIV  
XX infection. They may additionally be used as a contraceptive as HA  
XX mediates binding between the sperm and the oocyte. The peptides affords  
XX better inhibition of glycosaminoglycan-mediated activity than larger,  
XX less specific glycosaminoglycan inhibitors such as receptor antibodies,  
XX e.g., anti-CD4 antibodies, and are more specific than chemical  
XX inhibitors of glycosaminoglycans such as tunicamycin and H7. The peptides  
XX are significantly smaller than other inhibitors of  
XX glycosaminoglycan-mediated activity such as antibodies or soluble

CC glycosaminoglycan, allowing for improved oral and topical formulations,  
CC and are more cost effective to produce. Sequences AAB23282-823286  
CC represent the hyaluronic acid-binding peptide of the invention.  
XX  
SQ Sequence 12 AA;

Query Match 28.1%; Score 27; DB 21; Length 12;  
Best Local Similarity 50.0%; Pred. No. 5.5e+02;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
OY 3 HCNFNDVTTR 12  
DB 3 hmqfnaltrv 12

RESULT 25  
AAB81811  
ID AAB81811 standard; protein; 13 AA.  
XX  
AC AAB81811;  
XX  
DT 16-NOV-1990 (first entry)  
XX  
DE Sequence of peptide with immunological properties and structure of  
DE HIV-2 and SIV.1 glycoproteins.  
XX  
KW Antigen; immunogen; vaccine; diagnostic; AIDS.  
XX  
KW Immunodeficiency virus.  
XX  
OS  
XX  
PN WO8805440-A.  
XX  
PD 28-JUL-1988.  
XX  
PF 15-JAN-1988; 88WO-FR00025.  
XX  
PR 14-APR-1987; 87FR-0005398.  
XX  
PA (TNSP ) INST PASTEUR (ALIZ/).  
XX  
PI Allzon M, Montagnier L, Guetard D, Clavfl F, Sonigo P, Guyader M;  
PI Tollais P, Chakrabarti L, Desrosiers R;  
XX WPI; 1988-220290/31.  
XX

DR New peptide(s) with immunological properties of HIV-2 envelope protein -  
XX have the structure of simian immune deficiency virus proteins,  
XX useful in diagnosis and of vaccine components  
XX  
PS Claim 12; Page 44; 86pp; French.  
XX  
XX New peptides which have immunological properties in common with those of  
XX the peptide skeleton of the envelope protein of HIV-2 and also have a  
XX peptide structure in common with that of SIV-1 glycoprotein are claimed.  
XX Antigenic and immunogenic conjugates confg. the peptides and a kit to  
XX detect HIV-2 in biological fluids are new. The peptides are useful for in  
XX vitro diagnosis of HIV-2 infection and some of them can be used as  
XX components of immunogens and vaccines against HIV. Antibodies raised  
XX against them can be used for treatment of AIDS.

SQ Sequence 13 AA;

Query Match 28.1%; Score 27; DB 9; Length 13;  
Best Local Similarity 50.0%; Pred. No. 6e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 HCNFNDVT 10  
DB 3 hmqfnaltrv 10



Tue Apr 2 09:51:44 2002

us-09-020-393b-3\_copy\_42\_58.closed.rag

Page 13

Search completed: April 2, 2002, 09:24:40  
Job time: 109 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:23:26 ; Search time 90.24 Seconds  
(without alignments)  
52.307 Million cell updates/sec

Title: US-09-020-393B-3\_COPY\_42\_58  
Perfect score: 96  
Sequence: 1 FEHCNFNDVTTRLRENE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 359122

Minimum DB seq length: 0  
Maximum DB seq length: 17

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US060\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	44.8	15	22	US-09-835-752-1
2	38	39.6	17	20	US-09-607-158-9
3	33	34.4	15	6	US-08-287-535-6
4	32	33.3	16	1	PCT-US94-08453-11
5	32	33.3	16	6	US-08-232-513-11
6	32	33.3	16	12	US-08-824-238-11
7	32	33.3	16	12	US-08-824-808-11
8	32	33.3	16	14	US-09-017-259-11
9	32	33.3	17	1	PCT-US97-04143-16

10	32	33.3	17	13	US-08-928-074-16	Sequence 16, Appl
11	32	33.3	17	13	US-08-928-074-16	Sequence 16, Appl
12	32	33.3	17	13	US-08-928-074-16	Sequence 16, Appl
13	32	33.3	17	15	US-09-172-713-11	Sequence 11, Appl
14	30	31.2	8	15	US-09-187-859-1594	Sequence 1594, Ap
15	30	31.2	8	15	US-09-187-859-1594	Sequence 1594, Ap
16	30	31.2	8	22	US-09-839-542-1594	Sequence 1594, Ap
17	30	31.2	9	12	US-08-821-739-1	Sequence 1, Appl
18	30	31.2	9	15	US-09-187-859-1597	Sequence 1597, Ap
19	30	31.2	9	15	US-09-187-859-1597	Sequence 1597, Ap
20	30	31.2	9	22	US-09-839-542-1597	Sequence 1597, Ap
21	30	31.2	16	1	PCT-US96-19065-21	Sequence 21, Appl
22	29	30.2	13	24	US-06-299-378-28	Sequence 28, Appl
23	29	30.2	13	24	US-06-300-614-28	Sequence 28, Appl
24	29	30.2	14	15	US-09-169-404-23	Sequence 23, Appl
25	28	29.2	9	3	US-07-669-789-4	Sequence 4, Appl
26	28	29.2	13	13	US-08-908-371B-4	Sequence 4, Appl
27	28	29.2	13	21	US-09-765-101-34	Sequence 34, Appl
28	28	29.2	13	24	US-06-231-266-34	Sequence 34, Appl
29	27.5	28.6	10	1	PCT-US96-17882-133	Sequence 133, App
30	27.5	28.6	10	16	US-09-258-947-133	Sequence 133, App
31	27	28.1	12	8	US-08-433-108-20	Sequence 20, Appl
32	27	28.1	12	9	US-08-513-366-20	Sequence 20, Appl
33	27	28.1	12	9	US-08-513-366A-20	Sequence 20, Appl
34	27	28.1	12	10	US-08-685-357-12	Sequence 12, Appl
35	27	28.1	12	18	US-09-444-251-20	Sequence 20, Appl
36	27	28.1	12	19	US-09-532-709B-1	Sequence 1, Appl
37	27	28.1	12	20	US-09-685-010-72	Sequence 72, Appl
38	27	28.1	14	1	PCT-US00-30036A-112	Sequence 112, App
39	27	28.1	14	1	PCT-US01-11968-129	Sequence 729, App
40	27	28.1	14	14	US-09-009-843C-58	Sequence 58, Appl
41	27	28.1	14	14	US-09-009-843D-58	Sequence 58, Appl
42	27	28.1	14	22	US-09-833-245-729	Sequence 729, App
43	27	28.1	15	17	US-09-328-913A-15	Sequence 15, Appl
44	27	28.1	15	21	US-09-773-416-4	Sequence 4, Appl
45	27	28.1	15	22	US-09-819-308-18	Sequence 18, Appl
46	27	28.1	16	15	US-09-161-094-16	Sequence 16, Appl
47	27	28.1	16	15	US-09-161-094A-16	Sequence 16, Appl
48	27	28.1	17	10	US-08-622-338B-29	Sequence 29, Appl
49	27	28.1	17	10	US-08-622-338B-29	Sequence 29, Appl
50	27	28.1	17	10	US-08-622-338B-29	Sequence 29, Appl

#### ALIGNMENTS

RESULT 1  
US-09-835-752-1  
; Sequence 1, Application US/09835752  
; GENERAL INFORMATION:  
; APPLICANT: Halperin, Jose  
; TITLE OF INVENTION: Methods, Products and Treatments for Diabetes  
; FILE REFERENCE: H0498/7137(ERC)  
; CURRENT APPLICATION NUMBER: US/09/835,752  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: US 06/203,254  
; PRIOR FILING DATE: 2000-05-08  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-835-752-1

Query Match 44.8%; Score 43; DB 22; Length 15;  
Best Local Similarity 87.5%; Pred. No. 6.9;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FEHCNFND 8  
DB 7 FEHCNFND 14

RESULT 2  
US-09-607-158-9  
; Sequence 9, Application US/09607158  
; GENERAL INFORMATION:  
; APPLICANT: SAKAKAKA, MASAHIRO  
; APPLICANT: TANAKA, JUNYA  
; APPLICANT: SATO, KOHJI  
; APPLICANT: MORITA, FUMIO  
; APPLICANT: SADAMOTO, YASUTAKA  
; TITLE OF INVENTION: CYTOPROTECTIVE AGENTS COMPRISING PROSAPOSIN-RELATED  
; FILE REFERENCE: 49960  
; CURRENT APPLICATION NUMBER: US/09/607,158  
; CURRENT FILING DATE: 2000-06-29  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-607-158-9

Query Match 39.6%; Score 38; DB 20; Length 17;  
Best Local Similarity 60.0%; Pred. No. 53;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 EHCNFNDVTT 11  
Db 2 EHCSLNENIT 11

RESULT 3  
US-08-287-535-6  
; Sequence 6, Application US/08287535  
; GENERAL INFORMATION:  
; APPLICANT: Moe, Gregory R.  
; TITLE OF INVENTION: Conformationally-Stabilized Peptides  
; NUMBER OF INVENTION: and use thereof As Immunogens  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Connolly and Hutz  
; STREET: 1220 Market Street  
; CITY: Wilmington  
; STATE: Delaware  
; COUNTRY: U.S.A.  
; ZIP: 19899  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette  
; COMPUTER: IBM/PC or Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/287,535  
; FILING DATE: Not yet Assigned  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA: None  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 Amino Acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
; MOLECULE TYPE: Peptide  
US-08-287-535-6

Query Match 34.4%; Score 33; DB 6; Length 15;  
Best Local Similarity 45.5%; Pred. No. 3e+02;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
QY 4 CNFNDVTTLR 14

Db 1 CNFNEKLSQL 11  
RESULT 4  
PCT-US94-08453-11  
; Sequence 11, Application PC/TUS9408453  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbde, Martens, Olson & Bear  
; STREET: 620 Newport Center Drive, Sixteenth Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/08453  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/100,247  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: OBRLEN.002CPI  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: Internal  
PCT-US94-08453-11

Query Match 33.3%; Score 32; DB 1; Length 16;  
Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EHCNFNDVTT 11  
Db 2 EHCSLNENIT 11

RESULT 5  
US-08-232-513-11  
; Sequence 11, Application US/08232513  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbde, Martens, Olson & Bear  
; STREET: 620 Newport Center Drive, Sixteenth Floor  
; CITY: Newport Beach  
; STATE: CA

COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,513  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/100,247  
FILING DATE: 30-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelisen, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: O'BRIEN, 002CPI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ. ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
US-08-232-513-11

Query Match 33.3%; Score 32; DB 6; Length 16;  
Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EHCNFNDVTT 11  
| | | : | : |  
Db 2 EHCNSLNNIT 11

RESULT 6  
US-08-824-238-11  
Sequence 11, Application US/08824238  
GENERAL INFORMATION:  
APPLICANT: O'Brien, John S.  
TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides  
TITLE OF INVENTION: as Therapeutic Agents  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 620 Newport Center Drive, Sixteenth Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/824,238  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/100,247  
FILING DATE: 30-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelisen, Ned A.  
REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: O'BRIEN, 002CPI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ. ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
US-08-824-238-11

Query Match 33.3%; Score 32; DB 12; Length 16;  
Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EHCNFNDVTT 11  
| | | : | : |  
Db 2 EHCNSLNNIT 11

RESULT 7  
US-08-824-808-11  
Sequence 11, Application US/08824808  
GENERAL INFORMATION:  
APPLICANT: O'Brien, John S.  
TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides  
TITLE OF INVENTION: as Therapeutic Agents  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 620 Newport Center Drive, Sixteenth Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/824,808  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/100,247  
FILING DATE: 30-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelisen, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: O'BRIEN, 002CPI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ. ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
US-08-824-808-11

Query Match 33.3%; Score 32; DB 12; Length 16;  
Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EHCNFNDVTT 11  
Db 2 EHCNSLNNIT 11

RESULT 8  
US-09-017-259-11  
; Sequence 11, Application US/09017259  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides  
; TITLE OF INVENTION: as Therapeutic Agents  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 620 Newport Center Drive, Sixteenth Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/017,259  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/824,238  
; FILING DATE:  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: O'Brien, 002CP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; US-09-017-259-11  
  
Query Match 33.3%; Score 32; DB 14; Length 16;  
Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EHCNFNDVTT 11  
Db 2 EHCNSLNNIT 11  
  
RESULT 9  
PCT-US97-04143-16  
; Sequence 16, Application PC/TUS9704143  
; GENERAL INFORMATION:  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Methods of Alleviating Neuropathic Pain

;; TITLE OF INVENTION: Using Prosaposin-Derived Peptides  
;; NUMBER OF SEQUENCES: 21  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Campbell and Flores  
;; STREET: 4370 La Jolla Village Drive, Suite 700  
;; CITY: San Diego  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US97/04143  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: EP-UD 2474  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; PCT-US97-04143-16

Query Match 33.3%; Score 32; DB 1; Length 17;  
Best Local Similarity 50.0%; Pred. No. 5.1e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EHCNFNDVTT 11  
Db 2 EHCNSLNNIT 11

RESULT 10  
US-08-928-074-16  
; Sequence 16, Application US/08928074  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; TITLE OF INVENTION: METHODS FOR ALLEVIATING NEUROPATHIC PAIN  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,074  
; FILING DATE: 11-SEP-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/611,307  
; FILING DATE: 05-MAR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US97/04143  
; FILING DATE: 05-MAR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halle, Ph.D., Lisa A.  
; REGISTRATION NUMBER: 38,347

```

; REFERENCE/DOCKET NUMBER: 07256/024001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-928-074-16

Query Match          33.3%; Score 32; DB 13; Length 17;
Best Local Similarity 50.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EHCNENDVTT 11
   |||: | : |
Db 2 EHCSLNEKIT 11

RESULT 11
US-08-928-074-16
; Sequence 16, Application US/08928074A
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; TITLE OF INVENTION: METHODS FOR ALLEVIATING NEUROPATHIC PAIN
; FILE REFERENCE: 07256/024001
; CURRENT APPLICATION NUMBER: US/08/928,074A
; EARLIER FILING DATE: 1997-09-11
; EARLIER APPLICATION NUMBER: 08/611,307
; EARLIER FILING DATE: 1996-03-05
; EARLIER APPLICATION NUMBER: PCT/US97/04143
; EARLIER FILING DATE: 1996-03-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Artificial
US-08-928-074-16

Query Match          33.3%; Score 32; DB 13; Length 17;
Best Local Similarity 50.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EHCNENDVTT 11
   |||: | : |
Db 2 EHCSLNEKIT 11

RESULT 12
US-08-928-074-16
; Sequence 16, Application US/08928074B
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; TITLE OF INVENTION: METHODS FOR ALLEVIATING NEUROPATHIC PAIN
; FILE REFERENCE: 07256/024001
; CURRENT APPLICATION NUMBER: US/08/928,074B
; EARLIER FILING DATE: 1997-09-11
; EARLIER APPLICATION NUMBER: 08/611,307
; EARLIER FILING DATE: 1996-03-05
; EARLIER APPLICATION NUMBER: PCT/US97/04143
; EARLIER FILING DATE: 1996-03-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 17
```

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Artificial
US-08-928-074-16

Query Match          33.3%; Score 32; DB 13; Length 17;
Best Local Similarity 50.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EHCNENDVTT 11
   |||: | : |
Db 2 EHCSLNEKIT 11

RESULT 13
US-09-172-713-11
; Sequence 11, Application US/09172713
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/172,713
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/232,513
; FILING DATE: 21-APR-1994
; APPLICATION NUMBER: US 08/100,247
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 1643
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..17
; OTHER INFORMATION: /label= hepo
US-09-172-713-11

Query Match          33.3%; Score 32; DB 15; Length 17;
Best Local Similarity 50.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EHCNENDVTT 11
   |||: | : |
Db 2 EHCSLNEKIT 11
```

RESULT 14  
US-09-187-859-1594  
; Sequence 1594, Application US/09187859  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407C1  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1594  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
US-09-187-859-1594  
; OTHER INFORMATION: cadherin-8 cell adhesion recognition sequence

Query Match 31.2%; Score 30; DB 15; Length 8;  
Best Local Similarity 71.4%; Pred. No. 2.9e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 4 CINFNDVT 10  
Db 1 CQINDVT 7

RESULT 15  
US-09-187-859-1594  
; Sequence 1594, Application US/09187859A  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1594  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
US-09-187-859-1594  
; OTHER INFORMATION: cadherin-8 cell adhesion recognition sequence

Query Match 31.2%; Score 30; DB 15; Length 8;  
Best Local Similarity 71.4%; Pred. No. 2.9e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 4 CINFNDVT 10  
Db 1 CQINDVT 7

RESULT 16  
US-09-839-542-1594  
; Sequence 1594, Application US/09839542  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1594  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
US-09-839-542-1594  
; OTHER INFORMATION: cadherin-8 cell adhesion recognition sequence

Query Match 31.2%; Score 30; DB 22; Length 8;  
Best Local Similarity 71.4%; Pred. No. 2.9e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 4 CINFNDVT 10  
Db 1 CQINDVT 7

RESULT 17  
US-08-821-739-1  
; Sequence 1, Application US/08821739  
; GENERAL INFORMATION:  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Sette, Alessandro  
; TITLE OF INVENTION: HLA Binding Peptides and Their  
; NUMBER OF SEQUENCES: 149  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; APPLICATION DATA:  
; APPLICATION NUMBER: US/08/821,739  
; FILING DATE: 20-MAR-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/013,833  
; FILING DATE: 21-MAR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen Lauer  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 018623-0005081US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-821-739-1



Query Match 31.2% Score 30; DB 12; Length 9;  
Best Local Similarity 62.5%; Pred. No. 2.9e+06;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 HCNFNDVT 10  
| | | | |  
DB 1 HSNINDVT 8

RESULT 18  
US-09-187-859-1597

; Sequence 1597, Application US/09187859  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1597  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-8 cell adhesion recognition sequence  
US-09-187-859-1597

Query Match 31.2% Score 30; DB 15; Length 9;  
Best Local Similarity 71.4%; Pred. No. 2.9e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CNFNDVT 10  
| | | | |  
DB 1 CQINDVT 7

RESULT 19  
US-09-187-859-1597  
; Sequence 1597, Application US/09187859A  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1597  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-8 cell adhesion recognition sequence  
US-09-187-859-1597

Query Match 31.2% Score 30; DB 15; Length 9;  
Best Local Similarity 71.4%; Pred. No. 2.9e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CNFNDVT 10  
| | | | |  
DB 1 CQINDVT 7

RESULT 20  
US-09-839-542-1597

; Sequence 1597, Application US/09839542  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1597  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-8 cell adhesion recognition sequence  
US-09-839-542-1597

Query Match 31.2% Score 30; DB 22; Length 9;  
Best Local Similarity 71.4%; Pred. No. 2.9e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CNFNDVT 10  
| | | | |  
DB 1 CQINDVT 7

RESULT 21  
PCT-US96-19065-21  
; Sequence 21, Application PC/TUS9619065  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Metal-Regulated Transporters and Uses  
; NUMBER OF SEQUENCES: 21  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/19065  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/018,578  
; FILING DATE: 29-MAY-1996  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Silveri, Jean M.  
; REGISTRATION NUMBER: 39,030  
; REFERENCE/DOCKET NUMBER: DCI-099CPC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)227-5941  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: Internal

PCT-US96-19065-21

## Query Match

Best Local Similarity 31.2%; Score 30; DB 1; Length 16;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;QY 7 NDVTTRLRENE 17  
| | | | : | : |  
DB 3 NDVTLPIKEDD 13

## RESULT 22

US-60-299-378-28

; Sequence 28, Application US/60299378

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL POTASSIUM CHANNEL WITH HOMOLOGY

; TITLE OF INVENTION: Ether-a-90-90 FAMILY, HEAG2, EXPRESSED SPECIFICALLY IN THALAMUS,

; FILE REFERENCE: D0123

; CURRENT APPLICATION NUMBER: US/60/299,378

; CURRENT FILING DATE: 2001-06-19

; NUMBER OF SEQ ID NOS: 52

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 28

; LENGTH: 13

; TYPE: PRT

; ORGANISM: Homo sapiens

US-60-299-378-28

## Query Match

Best Local Similarity 30.2%; Score 29; DB 24; Length 13;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;QY 4 CNFNDVT 10  
| | | | : | : |  
DB 5 CTFKDT 11

## RESULT 23

US-60-300-614-28

; Sequence 28, Application US/60300614

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL POTASSIUM CHANNEL WITH HOMOLOGY

; TITLE OF INVENTION: Ether-a-90-90 FAMILY, HEAG2, EXPRESSED SPECIFICALLY IN THALAMUS,

; FILE REFERENCE: D0123 PSPI

; CURRENT APPLICATION NUMBER: US/60/300,614

; CURRENT FILING DATE: 2001-06-25

; NUMBER OF SEQ ID NOS: 52

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 28

; LENGTH: 13

; TYPE: PRT

; ORGANISM: Homo sapiens

US-60-300-614-28

## Query Match

Best Local Similarity 30.2%; Score 29; DB 24; Length 13;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;QY 4 CNFNDVT 10  
| | | | : | : |  
DB 5 CTFKDT 11

## RESULT 24

US-09-169-404-23

; Sequence 23, Application US/09169404

## GENERAL INFORMATION:

APPLICANT: Engvall, Eva

APPLICANT: Lelivo, Ilmo

TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin

TITLE OF INVENTION: Fragments and Uses Thereof

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESS: Campbell and Flores

STREET: 4370 La Jolla Village Drive, Suite 700

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/169,404

FILING DATE: Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/125,077

FILING DATE: 22-SEP-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US PCT/US 94/10730

FILING DATE: 21-SEP-1994

APPLICATION DATA:

APPLICATION NUMBER: US 07/472,319

FILING DATE: 30-JAN-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/919,951

FILING DATE: 27-JUL-1992

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-LA 9721

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 23:

LENGTH: 14 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-09-169-404-23

## Query Match

Best Local Similarity 30.2%; Score 29; DB 15; Length 14;  
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;QY 4 CNFNDVTTLREN 16  
| | | | : | : |  
DB 1 CSTVDITNOEN 13

## RESULT 25

US-07-669-789-4

; Sequence 4, Application US/07669789

; GENERAL INFORMATION:

; APPLICANT: Liu, Chung-Cheng

; TITLE OF INVENTION: Use of Purified Ubiquitin Hydrolase in Recovering Polypepti

; NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/669,789  
FILING DATE: 19910315  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/210,909  
FILING DATE: June 24, 1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 513D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/266-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
US-07-669-789-4

Query Match 29.2%; Score 28; DB 3; Length 9;  
Best Local Similarity 71.4%; Pred. NO. 2.9e+06;  
Matches 5; Conservative 2; Mismatches 0; Indels 0;  
Caps 0;  
QY 6 FNDVTR 12  
I:||||:  
Db 3 FDDVTTQ 9

Search completed: April 2, 2002, 09:26:37  
Job time: 191 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:23:56 ; Search time 15.54 Seconds  
(without alignments)  
78.189 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58  
Perfect score: 96  
Sequence: 1 FEHCNFNDVTTRLENE 17

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 348507 seqs, 71473746 residues

Total number of hits satisfying chosen parameters: 93745

Minimum DB seq length: 0  
Maximum DB seq length: 17

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database :

Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
8: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	31	32.3	15	6	US-09-989-919-123	Sequence 123, App
2	30	31.2	8	7	US-10-006-869-1594	Sequence 1594, App
3	30	31.2	9	7	US-10-006-869-1597	Sequence 1597, App
4	30	31.2	16	6	US-09-579-174-21	Sequence 21, App
5	29	30.2	11	6	US-09-754-831A-22	Sequence 22, App
6	28	29.2	10	6	US-09-573-822C-387	Sequence 387, App
7	28	29.2	13	1	PCT-US01-27692A-5	Sequence 5, App1
8	28	29.2	13	6	US-09-701-080A-5	Sequence 5, App1
9	28	29.2	13	6	US-09-701-080A-10	Sequence 10, App
10	27	28.1	12	6	US-09-965-738-260	Sequence 260, App
11	27	28.1	14	1	PCT-US02-08123-1760	Sequence 1760, App
12	27	28.1	16	7	US-10-038-612-16	Sequence 16, App
13	27	28.1	17	7	US-10-038-612-114	Sequence 114, App
14	26	27.1	12	4	US-08-816-454B-208	Sequence 208, App
15	26	27.1	12	6	US-09-921-822-33	Sequence 33, App
16	26	27.1	12	6	US-09-954-385-373	Sequence 373, App
17	26	27.1	12	7	US-10-084-813-422	Sequence 422, App
18	26	27.1	15	6	US-09-828-574-7	Sequence 7, App1
19	26	27.1	15	7	US-10-084-813-419	Sequence 419, App
20	25	26.0	7	7	US-10-006-869-1592	Sequence 1592, App
21	25	26.0	8	1	PCT-US01-27702A-63	Sequence 63, App
22	25	26.0	8	5	US-09-918-873-15	Sequence 15, App
23	25	26.0	10	6	US-09-572-270B-50	Sequence 50, App
24	25	26.0	10	6	US-09-572-270B-54	Sequence 54, App
25	26.0	10	6	US-09-572-270B-58	Sequence 58, App	

26	25	26.0	10	6	US-09-572-270B-62	Sequence 62, Appl
27	25	26.0	10	6	US-09-572-270B-66	Sequence 66, Appl
28	25	26.0	10	6	US-09-572-270B-70	Sequence 70, Appl
29	25	26.0	10	6	US-09-832-230-67	Sequence 67, Appl
30	25	26.0	10	6	US-09-572-404B-549	Sequence 549, Appl
31	25	26.0	10	6	US-09-572-404B-550	Sequence 550, Appl
32	25	26.0	10	6	US-09-572-404B-3662	Sequence 3662, Ap
33	25	26.0	10	6	US-09-572-404B-3663	Sequence 3663, Ap
34	25	26.0	10	6	US-09-832-230A-67	Sequence 67, Appl
35	25	26.0	10	7	US-10-083-768-67	Sequence 67, Appl
36	25	26.0	13	6	US-09-952-677-8	Sequence 8, Appl
37	25	26.0	14	7	US-10-005-549-28	Sequence 28, Appl
38	25	26.0	15	6	US-09-443-986A-1	Sequence 1, Appl
39	25	26.0	15	6	US-09-711-161-14	Sequence 14, Appl
40	25	26.0	17	6	US-09-834-784-848	Sequence 848, Appl
41	25	26.0	17	6	US-09-350-641C-848	Sequence 848, Appl
42	24	25.0	9	7	US-10-084-813-423	Sequence 423, Appl
43	24	25.0	11	7	US-10-020-269-26	Sequence 26, Appl
44	24	25.0	12	7	US-10-084-813-420	Sequence 420, Appl
45	24	25.0	12	7	US-10-084-813-421	Sequence 421, Appl
46	24	25.0	13	6	US-09-419-901A-4	Sequence 4, Appl
47	24	25.0	15	7	US-10-038-407-13	Sequence 13, Appl
48	24	25.0	15	7	US-10-084-813-417	Sequence 417, Appl
49	24	25.0	15	7	US-10-084-813-418	Sequence 418, Appl
50	24	25.0	17	6	US-09-350-641C-1523	Sequence 1523, Ap

#### ALIGNMENTS

```
RESULT 1
US-09-989-919-123
; Sequence 123, Application US/09989919
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Requion, Hervé
; APPLICANT: Pluta, Jason
; APPLICANT: Ghosh, Malavika
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and
; FILE REFERENCE: DEX-0289
; CURRENT APPLICATION NUMBER: US/09/989,919
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,505
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 123
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-919-123

Query Match          32.3%; Score 31; DB 6; Length 15;
Best Local Similarity 30.0%; Pred. No. 56;
Matches 3; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY      1 FEHCNFNDVT 10
       :::::|:::
Db       4 YKHCHMNNLS 13

RESULT 2
US-10-006-869-1594
; Sequence 1594, Application US/10006869
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; CADHERIN-MEDIATED FUNCTIONS
```

FILE REFERENCE: 100086.407C7  
CURRENT APPLICATION NUMBER: US/10/006,869  
CURRENT FILING DATE: 2001-12-03  
NUMBER OF SEQ ID NOS: 4052  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1594  
LENGTH: 8  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Representative cyclic modulating agent based on  
US-10-006-869-1594

Query Match  
Best Local Similarity 31.2%; Score 30; DB 7; Length 8;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CNFNQVT 10  
DB 1 COINDVT 7

RESULT 3  
US-10-006-869-1597  
Sequence 1597, Application US/10006869  
GENERAL INFORMATION:  
APPLICANT: Blaschuk, Orest W.  
APPLICANT: Symonds, James Matthew  
APPLICANT: Gour, Barbara J.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
FILE REFERENCE: 100086.407C7  
CURRENT APPLICATION NUMBER: US/10/006,869  
CURRENT FILING DATE: 2001-12-03  
NUMBER OF SEQ ID NOS: 4052  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1597  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Representative cyclic modulating agent based on  
US-10-006-869-1597

Query Match  
Best Local Similarity 31.2%; Score 30; DB 7; Length 9;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CNFNQVT 10  
DB 1 COINDVT 7

RESULT 4  
US-09-579-174-21  
Sequence 21, Application US/09579174  
GENERAL INFORMATION:  
APPLICANT: Gueriot, Mary Lou et al.  
TITLE OF INVENTION: METAL-REGULATED TRANSPORTERS AND USES THEREFOR  
FILE REFERENCE: DCI-099CPDV  
CURRENT APPLICATION NUMBER: US/09/579,174  
CURRENT FILING DATE: 2000-05-25  
PRIOR APPLICATION NUMBER: 09/107,858  
PRIOR FILING DATE: 1998-06-30  
PRIOR APPLICATION NUMBER: 08/758,621  
PRIOR FILING DATE: 1996-11-27  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 21

LENGTH: 16  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-579-174-21

Query Match  
Best Local Similarity 31.2%; Score 30; DB 6; Length 16;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 7 NDVTRLRENE 17  
DB 3 NDVTRLRKED 13

RESULT 5  
US-09-754-831A-22  
Sequence 22, Application US/09754831A  
GENERAL INFORMATION:  
APPLICANT: Oppermann, Herman  
APPLICANT: Kuberassampath, Thangavel  
APPLICANT: Rueger, David  
TITLE OF INVENTION: Osteogenic Devices  
FILE REFERENCE: STR-008CN  
CURRENT APPLICATION NUMBER: US/09/754,831A  
CURRENT FILING DATE: 2001-01-03  
PRIOR APPLICATION NUMBER: US 08/375,901  
PRIOR FILING DATE: 1995-01-20  
PRIOR APPLICATION NUMBER: US 08/145,812  
PRIOR FILING DATE: 1993-11-01  
PRIOR APPLICATION NUMBER: US 07/995,345  
PRIOR FILING DATE: 1992-12-22  
PRIOR APPLICATION NUMBER: US 07/315,342  
PRIOR FILING DATE: 1989-02-23  
PRIOR APPLICATION NUMBER: US 07/232,630  
PRIOR FILING DATE: 1988-08-15  
PRIOR APPLICATION NUMBER: US 07/179,406  
PRIOR FILING DATE: 1988-04-08  
NUMBER OF SEQ ID NOS: 72  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 22  
LENGTH: 11  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: peptide fragment  
US-09-754-831A-22

Query Match  
Best Local Similarity 30.2%; Score 29; DB 6; Length 11;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 PEHCNEND 8  
DB 1 FLHCQFSE 8

RESULT 6  
US-09-573-822C-387  
Sequence 387, Application US/09573822C  
GENERAL INFORMATION:  
APPLICANT: Proteom Ltd  
TITLE OF INVENTION: Complementary peptide ligands generated from microbial genome  
FILE REFERENCE: Microbe patent  
CURRENT APPLICATION NUMBER: US/09/573,822C  
CURRENT FILING DATE: 2000-05-18  
NUMBER OF SEQ ID NOS: 804  
SOFTWARE: ProCPatent version 1.0  
SEQ ID NO 387  
LENGTH: 10  
TYPE: PRT

ORGANISM: mycoplasma genitalium  
FEATURE:  
OTHER INFORMATION: Sequence located in MG121 at 123-132 and may interact with Sequen  
US-09-573-822C-387

Query Match 29.2%; Score 28; DB 6; Length 10;  
Best Local Similarity 55.6%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 6 FNDVTRLR 14  
|:|:|:|:  
Db 1 FSDTTLKRV 9

RESULT 7  
PCT-US01-27692A-5  
Sequence 5, Application PC/TUS0127692A  
GENERAL INFORMATION:  
APPLICANT: Board of Regents, The University of Texas System  
TITLE OF INVENTION: Human and Mouse Targeting Peptides Identified by Phage Display  
FILE REFERENCE: 005774.P003PCT  
CURRENT APPLICATION NUMBER: PCT/US01/27692A  
CURRENT FILING DATE: 2001-09-07  
NUMBER OF SEQ ID NOS: 251  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5  
LENGTH: 13  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: (1):(13)  
OTHER INFORMATION: synthetic construct  
PCT-US01-27692A-5

Query Match 29.2%; Score 28; DB 1; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHC 4  
|:|:|:|:  
Db 10 FEHC 13

RESULT 8  
US-09-701-080A-5  
Sequence 5, Application US/09701080A  
GENERAL INFORMATION:  
APPLICANT: O'CONNOR, MARK J.  
APPLICANT: ZIMMERMAN, HOLGER  
TITLE OF INVENTION: POLYPEPTIDES FROM CREB BINDING PROTEIN AND RELATED PROTEIN P300 F  
FILE REFERENCE: 117-328  
CURRENT APPLICATION NUMBER: US/09/701,080A  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: GB 9811303.8  
PRIOR FILING DATE: 1998-05-26  
PRIOR APPLICATION NUMBER: GB 9900157.0  
PRIOR FILING DATE: 1999-01-05  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 13  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:derived from E2F  
US-09-701-080A-5

Query Match 29.2%; Score 28; DB 6; Length 13;  
Best Local Similarity 57.1%; Pred. No. 1.5e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 CNFNDVTR 10  
|:|:|:|:  
Db 3 CDFGDLTR 9

RESULT 9  
US-09-701-080A-10  
Sequence 10, Application US/09701080A  
GENERAL INFORMATION:  
APPLICANT: O'CONNOR, MARK J.  
APPLICANT: ZIMMERMAN, HOLGER  
TITLE OF INVENTION: POLYPEPTIDES FROM CREB BINDING PROTEIN AND RELATED PROTEIN P30  
FILE REFERENCE: 117-328  
CURRENT APPLICATION NUMBER: US/09/701,080A  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: GB 9811303.8  
PRIOR FILING DATE: 1998-05-26  
PRIOR APPLICATION NUMBER: GB 9900157.0  
PRIOR FILING DATE: 1999-01-05  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 10  
LENGTH: 13  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:derived from E2F  
US-09-701-080A-10

Query Match 29.2%; Score 28; DB 6; Length 13;  
Best Local Similarity 57.1%; Pred. No. 1.5e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 CNFNDVTR 10  
|:|:|:|:  
Db 3 CDFGDLTR 9

RESULT 10  
US-09-965-738-260  
Sequence 260, Application US/09965738  
GENERAL INFORMATION:  
APPLICANT: O'Brien, Timothy  
TITLE OF INVENTION: Repeat Sequences of the CA125 Gene and Their Use for Diagnosti  
FILE REFERENCE: 40715-258841  
CURRENT APPLICATION NUMBER: US/09/965,738  
CURRENT FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: US 60/284,175  
PRIOR FILING DATE: 2001-04-17  
NUMBER OF SEQ ID NOS: 306  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 260  
LENGTH: 12  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-965-738-260

Query Match 28.1%; Score 27; DB 6; Length 12;  
Best Local Similarity 45.5%; Pred. No. 2.1e+02;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 FEHCNFDVTR 11  
|:|:|:|:  
Db 1 FTHRNVPVITS 11

RESULT 11  
PCT-US02-08123-1760  
; Sequence 1760, Application PC/TUS0208123  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS904PCT  
; CURRENT APPLICATION NUMBER: PCT/US02/08123  
; CURRENT FILING DATE: 2002-03-19  
; PRIOR APPLICATION NUMBER: US 60/331,287  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/306,171  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: US 60/277,340  
; PRIOR FILING DATE: 2001-03-21  
; NUMBER OF SEQ ID NOS: 2048  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1760  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-08123-1760

Query Match 28.1%; Score 27; DB 1; Length 14;  
Best Local Similarity 71.4%; Pred. No. 2.9e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 NDVTTRL 13  
Db 5 NDVTAKL 11

RESULT 12  
US-10-038-612-16  
; Sequence 16, Application US/10038612  
; GENERAL INFORMATION:  
; APPLICANT: Ben-Sasson, Shmuel A.  
; TITLE OF INVENTION: Short Peptides Which Selectively  
; FILE REFERENCE: 1242.1029-000 (CMCC-679)  
; CURRENT APPLICATION NUMBER: US/10/038,612  
; CURRENT FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 09/161,094  
; PRIOR FILING DATE: 1998-09-25  
; NUMBER OF SEQ ID NOS: 172  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: unknown  
; FEATURE:  
; OTHER INFORMATION: CK IIA  
US-10-038-612-16

Query Match 28.1%; Score 27; DB 7; Length 16;  
Best Local Similarity 62.5%; Pred. No. 2.9e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 FEHCNFND 8  
Db 1 FEHVNTMD 8

RESULT 13  
US-10-038-612-114  
; Sequence 114, Application US/10038612  
; GENERAL INFORMATION:  
; APPLICANT: Ben-Sasson, Shmuel A.  
; TITLE OF INVENTION: Short Peptides Which Selectively  
; TITLE OF INVENTION: Modulate the Activity of Protein Kinases

FILE REFERENCE: 1242.1029-000 (CMCC-679)  
; CURRENT APPLICATION NUMBER: US/10/038,612  
; CURRENT FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 09/161,094  
; PRIOR FILING DATE: 1998-09-25  
; NUMBER OF SEQ ID NOS: 172  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 114  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: MYRISTATE  
; LOCATION: (1)-(10)  
; NAME/KEY: AMIDATION  
; LOCATION: (0)...(17)  
; OTHER INFORMATION: CK IIA  
US-10-038-612-114

Query Match 28.1%; Score 27; DB 7; Length 17;  
Best Local Similarity 62.5%; Pred. No. 3.1e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 FEHCNFND 8  
Db 2 FEHVNTMD 9

RESULT 14  
US-08-816-454B-208  
; Sequence 208, Application US/08816454B  
; GENERAL INFORMATION:  
; APPLICANT: SLINGLUFF, Craig L.  
; HUNT, Donald F.  
; SHABANOWITZ, Jeffrey  
; COX, Andrea L.  
; ENGELHARD, Victor H.  
; KITTELSEN, David  
; SKIPPER, Jonathan  
; HENDRIKSON, Ronald C.  
; TITLE OF INVENTION: PEPTIDES RECOGNIZED BY MELANOMA-SPECIFIC  
; A1-, A2, A3-RESTRICTED CYTOTOXIC LYMPHOCYTES AND USES  
; NUMBER OF SEQUENCES: 294  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NIEMARK, P.L.L.C.  
; STREET: 624 Ninth Street N.W., Ste. 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/816,454B  
; FILING DATE: 12-Mar-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/027,627  
; FILING DATE: 04-OCT-1996  
; APPLICATION NUMBER: US 60/013,972  
; FILING DATE: 19-MAR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: COOPER, Iver P.  
; REGISTRATION NUMBER: 28,005  
; REFERENCE/DOCKET NUMBER: SLINGLUFF-2/3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 628-5197  
; TELEFAX: (202) 737-3528



9  
; INFORMATION FOR SEQ ID NO: 208:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 208:  
US-08-816-454B-208

Query Match 27.1%; Score 26; DB 4; Length 12;  
Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 FNDVTT 11  
111111  
DB 6 FNDINT 11

## RESULT 15

US-09-921-822-33  
; Sequence 33, Application US/09921822

; GENERAL INFORMATION:

; APPLICANT: Schatz, Peter J.

; Miller, Jeff F.

; Cull, Millard G.

; Stemmer, Willem P.C.

; Gates, Christian M.

; TITLE OF INVENTION: Peptide Library and Screening Method

; NUMBER OF SEQUENCES: 162

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: William M. Smitch

; STREET: One Market Plaza, Stewart Tower, Suite 2000

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94105

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/921,822

; FILING DATE: 02-Aug-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/504,787

; FILING DATE: <Unknown>

; APPLICATION NUMBER: US 08/290,641

; FILING DATE: 15-Aug-1994

; APPLICATION NUMBER: US 07/963,321

; FILING DATE: 15-Oct-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Smitch, William M.

; REGISTRATION NUMBER: 30,223

; REFERENCE/DOCKET NUMBER: 16528J-001240US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-326-2400

; TELEFAX: 415-326-2422

; INFORMATION FOR SEQ ID NO: 33:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 12 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; IMMEDIATE SOURCE:

; CLONE: 7 4 1.1

; SEQUENCE DESCRIPTION: SEQ ID NO: 33:

; US-09-921-822-33

Query Match 27.1%; Score 26; DB 6; Length 12;  
Best Local Similarity 55.6%; Pred. No. 3.1e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 VTTRLRENE 17  
11111111  
DB 1 VTQVRVSNK 9

## RESULT 16

US-09-954-385-373

; Sequence 373, Application US/09954385

; GENERAL INFORMATION:

; APPLICANT: Aehle, Wolfgang

; APPLICANT: Baldwin, Toby L.

; APPLICANT: Van Gastel, Franciscus J.C.

; APPLICANT: Janssen, Giselle G.

; APPLICANT: Murray, Christopher J.

; APPLICANT: Wang, Huaming

; APPLICANT: Winetzk, Deborah S.

; TITLE OF INVENTION: Binding Phenol Oxidizing Enzyme-peptide

; FILE REFERENCE: Complexes

; CURRENT APPLICATION NUMBER: US/09/954,385

; CURRENT FILING DATE: 2001-09-12

; NUMBER OF SEQ ID NOS: 433

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 373

; LENGTH: 12

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: binding peptide

US-09-954-385-373

Query Match 27.1%; Score 26; DB 6; Length 12;  
Best Local Similarity 50.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 5 NFNDVTTT 12  
11111111  
DB 5 NFSNVSTK 12

## RESULT 17

US-10-084-813-422

; Sequence 422, Application US/10084813

; GENERAL INFORMATION:

; APPLICANT: SAXINGER, CARL

; TITLE OF INVENTION: POLYPEPTIDES THAT BIND HIV GP120 AND RELATED NUCLEIC

; FILE REFERENCE: 215875

; CURRENT APPLICATION NUMBER: US/10/084,813

; CURRENT FILING DATE: 2002-02-27

; PRIOR APPLICATION NUMBER: PCT/US00/23505

; PRIOR FILING DATE: 2000-08-25

; PRIOR APPLICATION NUMBER: US 60/151,270

; PRIOR FILING DATE: 1999-08-27

; NUMBER OF SEQ ID NOS: 1242

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 422

; LENGTH: 12

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: binding peptide

US-10-084-813-422

Query Match 27.1%; Score 26; DB 7; Length 12;  
Best Local Similarity 50.0%; Pred. No. 3.1e+02;

RESULT 20

RESULT 22  
 US-09-918-873-15  
 ; Sequence 15, Application US/09918873  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Goldsmith, Elizabeth J.  
 ; APPLICANT: Radha, Akella  
 ; APPLICANT: Gaylor, Richard B.  
 ; TITLE OF INVENTION: CHIMERIZING PROTEIN KINASES FOR DRUG  
 ; TITLE OF INVENTION: DISCOVERY  
 ; FILE REFERENCE: A33864 090495 0232

```
; CURRENT APPLICATION NUMBER: US/09/918,873
; CURRENT FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 15
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Homo sapiens
; DATABASE ACCESSION NUMBER: 18158777
; DATABASE ENTRY DATE: 2001-01-29
US-09-918-873-15
```

```
Query Match          26.0%; Score 25; DB 5; Length 8;
Best Local Similarity 75.0%; Pred. No. 2.9e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2 EHCN 5
    : 111
DB 2 DHCN 5
```

```
RESULT 23
US-09-572-270B-50
; Sequence 50, Application US/09572270B
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from plant genomes
; FILE REFERENCE: Plant patent
; CURRENT APPLICATION NUMBER: US/09/572,270B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 1144
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 50
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
; FEATURE:
; OTHER INFORMATION: Sequence located in CADI. at 36-45 and may interact with Sequence
US-09-572-270B-50
```

```
Query Match          26.0%; Score 25; DB 6; Length 10;
Best Local Similarity 57.1%; Pred. No. 3.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 7 NDVTTRL 13
    : 1111 :
DB 3 NDVTVKI 9
```

```
RESULT 24
US-09-572-270B-54
; Sequence 54, Application US/09572270B
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from plant genomes
; FILE REFERENCE: Plant patent
; CURRENT APPLICATION NUMBER: US/09/572,270B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 1144
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 54
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
; FEATURE:
; OTHER INFORMATION: Sequence located in CADI. at 36-45 and may interact with Sequence
US-09-572-270B-54
```

```
Query Match          26.0%; Score 25; DB 6; Length 10;
Best Local Similarity 57.1%; Pred. No. 3.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 7 NDVTTRL 13
    : 1111 :
DB 3 NDVTVKI 9
```

```
RESULT 25
US-09-572-270B-58
; Sequence 58, Application US/09572270B
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from plant genomes
; FILE REFERENCE: Plant patent
; CURRENT APPLICATION NUMBER: US/09/572,270B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 1144
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 58
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
; FEATURE:
; OTHER INFORMATION: Sequence located in CADI. at 36-45 and may interact with Sequence
US-09-572-270B-58
```

```
Query Match          26.0%; Score 25; DB 6; Length 10;
Best Local Similarity 57.1%; Pred. No. 3.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 7 NDVTTRL 13
    : 1111 :
DB 3 NDVTVKI 9
```

```
Search completed: April 2, 2002, 09:26:59
Job time: 183 sec
```

Tue Apr 2 09:51:50 2002

us-09-020-393b-3\_copy-42\_58.closed.ram

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:22:51 ; Search time 12.51 Seconds  
(without alignments)  
30.580 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58  
Perfect score: 96  
Sequence: 1 FEHCNFNDVTTRLRENE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues  
Total number of hits satisfying chosen parameters: 100354

Minimum DB seq length: 0  
Maximum DB seq length: 17

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database : Issued Patents\_AA:\*

1: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	33.3	17	1	US-08-232-513A-11
2	32	33.3	17	4	US-09-231-159-16
3	32	33.3	17	4	US-08-511-307-16
4	30	31.2	16	2	US-08-758-621-21
5	30	31.2	16	4	US-09-107-858-21
6	29	30.2	11	6	5496552-8
7	29	30.2	14	2	US-08-460-309-23
8	29	30.2	14	2	US-08-125-077-23
9	29	30.2	14	6	544158-4
10	28	29.2	15	1	US-08-522-326-6
11	27.5	28.6	10	2	US-08-556-597-133
12	27	28.1	12	2	US-08-685-357B-12
13	27	28.1	12	3	US-08-952-568-20
14	27	28.1	16	6	5194585-1
15	27	28.1	16	6	5194585-3
16	26.5	27.6	15	1	US-08-218-025A-105
17	26	27.1	11	4	US-09-224-785-22
18	26	27.1	11	4	US-09-224-785-25
19	26	27.1	12	1	US-07-778-233B-33
20	26	27.1	12	1	US-07-963-321-33
21	26	27.1	12	1	US-08-290-641-33
22	26	27.1	12	1	US-08-548-540-33
23	26	27.1	12	5	PCT-US96-09809-33
24	26	27.1	14	2	US-08-413-708B-5
25	26	27.1	14	3	US-09-192-048-24
26	26	27.1	16	4	US-09-164-186-5
27	26	27.1	17	4	US-08-706-344C-15

28	25.5	26.6	12	2	US-08-468-812-14	Sequence 14, Appl
29	25.5	26.6	12	4	US-08-590-563-22	Sequence 22, Appl
30	25	26.0	6	1	US-08-375-911A-5	Sequence 5, Appl1
31	25	26.0	9	4	US-09-502-600-69	Sequence 69, Appl
32	25	26.0	10	2	US-08-764-640-67	Sequence 67, Appl
33	25	26.0	10	3	US-08-973-225-67	Sequence 67, Appl
34	25	26.0	10	3	US-09-244-298A-67	Sequence 67, Appl
35	25	26.0	10	4	US-09-516-704-67	Sequence 67, Appl
36	25	26.0	11	4	US-09-224-785-29	Sequence 29, Appl
37	25	26.0	13	4	US-08-836-567-17	Sequence 17, Appl
38	25	26.0	13	4	US-09-196-390-8	Sequence 8, Appl1
39	25	26.0	15	1	US-08-221-583-17	Sequence 17, Appl
40	25	26.0	15	1	US-08-221-583-18	Sequence 18, Appl
41	25	26.0	15	1	US-08-221-583-19	Sequence 19, Appl
42	25	26.0	15	1	US-08-310-341B-1	Sequence 1, Appl1
43	25	26.0	15	5	PCT-US95-04018-17	Sequence 17, Appl
44	25	26.0	15	5	PCT-US95-04018-18	Sequence 18, Appl
45	25	26.0	15	5	PCT-US95-04018-19	Sequence 19, Appl
46	25	26.0	17	5	PCT-US95-05657-1	Sequence 1, Appl1
47	25	26.0	17	4	US-09-082-279B-848	Sequence 848, App
48	24	25.0	8	2	US-08-672-610A-53	Sequence 53, Appl
49	24	25.0	8	5	PCT-US91-02166-14	Sequence 14, Appl
50	24	25.0	9	4	US-09-518-046-74	Sequence 74, Appl

## ALIGNMENTS

RESULT 1  
US-08-232-513A-11  
Sequence 11, Application US/08232513A  
Patent No. 5700909  
GENERAL INFORMATION:  
APPLICANT: O'Brien, John S.  
TITLE OF INVENTION: Protoposin and Cytokine-Derived Peptides  
TITLE OF INVENTION: as Therapeutic Agents  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,513A  
FILING DATE: 21-Apr-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/100,247  
FILING DATE: 30-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-UD 1643  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..17  
OTHER INFORMATION: /label= hepo

US-08-232-513A-11

Query Match 33.3%; Score 32; DB 1; Length 17;  
Best Local Similarity 50.0%; Pred. No. 28;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EHCHNENDVTT 11  
Db 2 EHCHNENDVTT 11

RESULT 2

US-09-231-159-16  
; Sequence 16, Application US/092311159  
; Patent No. 6268347  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; TITLE OF INVENTION: Methods of Alleviating Neuropathic Pain  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/231,159  
; FILING DATE:  
; CLASSIFICATION:  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/231,159  
; FILING DATE: 05-MAR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-09-231-159-16

Query Match 33.3%; Score 32; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 28;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EHCHNENDVTT 11  
Db 2 EHCHNENDVTT 11

RESULT 3

US-08-611-307-16  
; Sequence 16, Application US/08611307  
; Patent No. 6271196  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; TITLE OF INVENTION: Methods of Alleviating Neuropathic Pain  
; NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/611,307  
FILING DATE: 05-MAR-1996  
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-UD 1928  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-8949  
TELEFAX: (619) 535-9001

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-611-307-16

Query Match 33.3%; Score 32; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 28;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EHCHNENDVTT 11  
Db 2 EHCHNENDVTT 11

RESULT 4

US-08-758-621-21  
; Sequence 21, Application US/08758621  
; Patent No. 5846821  
; GENERAL INFORMATION:  
; APPLICANT: Gueriot, Mary Lou, and Eide, David J.  
; TITLE OF INVENTION: Metal-Regulated Transporters and Uses Therefor  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/758,621  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/018,578  
; FILING DATE: 29-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Silveri, Jean M.  
; REGISTRATION NUMBER: 39,030  
; REFERENCE/DOCKET NUMBER: DCI-099CP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)227-5941

; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-08-758-621-21

Query Match 31.2%; Score 30; DB 2; Length 16;  
Best Local Similarity 45.5%; Pred. No. 57;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 7 NDVTTRLRENE 17  
| | | | : : : :  
Db 3 NDVTLPKEDD 13

RESULT 5  
US-09-107-858-21  
; Sequence 21, Application US/09107858  
; Patent No. 6162900  
; GENERAL INFORMATION:  
; APPLICANT: Gueriot, Mary Lou et al.  
; TITLE OF INVENTION: METAL-REGULATED TRANSPORTERS AND USES THEREFOR  
; FILE REFERENCE: DCI-099CPDV  
; CURRENT APPLICATION NUMBER: US/09/107,858  
; EARLIER FILING DATE: 1998-06-30  
; EARLIER APPLICATION NUMBER: 08/758,621  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 21  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-107-858-21

Query Match 31.2%; Score 30; DB 4; Length 16;  
Best Local Similarity 45.5%; Pred. No. 57;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 7 NDVTTRLRENE 17  
| | | | : : : :  
Db 3 NDVTLPKEDD 13

RESULT 6  
5496552-8  
; Patent No. 5496552  
; APPLICANT: KUBERASAMPATH, THANGAVEL, RUEGER, DAVID C.  
; TITLE OF INVENTION: OSTEOGENIC DEVICES  
; NUMBER OF SEQUENCES: 25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/268,252  
; FILING DATE: 29-JUN-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 103,604  
; FILING DATE: 06-AUG-1993  
; APPLICATION NUMBER: 827,052  
; FILING DATE: 28-JAN-1992  
; APPLICATION NUMBER: 579,865  
; FILING DATE: 07-SEP-1990  
; APPLICATION NUMBER: 179,406  
; FILING DATE: 08-APR-1988  
; SEQ ID NO: 8  
; LENGTH: 11  
5496552-8

Query Match 30.2%; Score 29; DB 6; Length 11;

Best Local Similarity 50.0%; Pred. No. 55;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 FEHCNFD 8  
| | | : : :  
Db 1 FLHCDFSE 8

RESULT 7  
US-08-460-309-23  
; Sequence 23, Application US/08460309  
; Patent No. 5837496  
; GENERAL INFORMATION:  
; APPLICANT: Leivo, Ilmo  
; TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/460,309  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/125,077  
; FILING DATE: 22-SEP-1993  
; APPLICATION NUMBER: US PCT/US 94/10730  
; FILING DATE: 21-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/472,319  
; FILING DATE: 30-JAN-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/919,951  
; FILING DATE: 27-JUL-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 9721  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-460-309-23

Query Match 30.2%; Score 29; DB 2; Length 14;  
Best Local Similarity 38.5%; Pred. No. 72;  
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 4 CNFNDVTTRLREN 16  
| : | : | : | :  
Db 1 CSIVDITNDEEN 13

RESULT 8  
US-08-125-077-23  
; Sequence 23, Application US/08125077  
; Patent No. 5872231





OY 1 EHCNFDVT 10  
| | | | |  
Db 4 FROCFENPIS 13

## RESULT 11

US-08-556-597-133  
; Sequence 133, Application US/08556597  
; Patent No. 5877155  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Lyle, Vicki A.  
; TITLE OF INVENTION: MIMOTOPES AND ANTI-MIMOTOPES OF  
; TITLE OF INVENTION: HUMAN PLATELET GLYCOPROTEIN Ib/IX  
; NUMBER OF SEQUENCES: 173  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/556,597  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/406,330  
; FILING DATE: 17-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Timian, Susan J.  
; REGISTRATION NUMBER: 34,103  
; REFERENCE/DOCKET NUMBER: 20884/101  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1636  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 133:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-556-597-133

Query Match 28.6%; Score 27.5; DB 2; Length 10;  
Best Local Similarity 46.2%; Pred. No. 87;  
Matches 6; Conservative 3; Mismatches 1; Indels 3; Gaps 1;

OY 2 EHCNFDVT 14  
| | | | |  
Db 1 KHCS---IHTRLR 10

## RESULT 12

US-08-685-357B-12  
; Sequence 12, Application US/08685357B  
; Patent No. 5854216  
; GENERAL INFORMATION:  
; APPLICANT: GAUDREAU, Pierrette  
; TITLE OF INVENTION: MARKER FOR GROWTH HORMONE-RELEASING  
; TITLE OF INVENTION: FACTOR RECEPTORS  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kevin M. Farrell, P.C.  
; STREET: 12 Riverwood Drive - P.O. Box 999  
; CITY: York Harbor

STATE: ME  
COUNTRY: U.S.A.  
ZIP: 03911  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/685,357B  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/312,244  
; FILING DATE: 23-SEP-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FARRELL, Kevin M.  
; REGISTRATION NUMBER: 35,505  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (207) 363-0558  
; TELEFAX: (207) 363-0528  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; US-08-685-357B-12

Query Match 28.1%; Score 27; DB 2; Length 12;  
Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 8 DVTRLRENE 17  
| | | | |  
Db 1 DFTQLRDE 10

RESULT 13  
US-08-952-568-20  
; Sequence 20, Application US/08952568  
; Patent No. 6037321  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Vasoactive Intestinal Peptide  
; NUMBER OF SEQUENCES: 43  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/952,568  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/CA96/00280  
; FILING DATE:  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: Internal  
; US-08-952-568-20

Query Match 28.1%; Score 27; DB 3; Length 12;  
Best Local Similarity 60.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 FNDVTRLRLRE 15  
| | | | |  
DB 1 FTDVTRLRLK 10

RESULT 14  
5194585-1  
; Patent No. 5194585  
; APPLICANT: PAUL, SUDHIR; POWELL, MICHAEL J.  
; MASSEY, RICHARD J.  
; TITLE OF INVENTION: INHIBITORS OF CATALYTIC ANTIBODIES  
; NUMBER OF SEQUENCES: 13  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/486,594  
; FILING DATE: 28-FEB-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 343,081  
; FILING DATE: 25-APR-1989  
; SEQ ID NO:1:  
; LENGTH: 16  
5194585-1

Query Match 28.1%; Score 27; DB 6; Length 16;  
Best Local Similarity 60.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 FNDVTRLRLRE 15  
| | | | |  
DB 6 FTDVTRLRLK 15

RESULT 15  
5194585-3  
; Patent No. 5194585  
; APPLICANT: PAUL, SUDHIR; POWELL, MICHAEL J.  
; MASSEY, RICHARD J.  
; TITLE OF INVENTION: INHIBITORS OF CATALYTIC ANTIBODIES  
; NUMBER OF SEQUENCES: 13  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/486,594  
; FILING DATE: 28-FEB-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 343,081  
; FILING DATE: 25-APR-1989  
; SEQ ID NO:3:  
; LENGTH: 16  
5194585-3

Query Match 28.1%; Score 27; DB 6; Length 16;  
Best Local Similarity 60.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 FNDVTRLRLRE 15  
| | | | |  
DB 6 FTDVTRLRLK 15

RESULT 16  
US-08-218-025A-105  
; Sequence 105; Application US/08218025A  
; Patent No. 5556744  
; GENERAL INFORMATION:  
; APPLICANT: Weiner, David B.  
; APPLICANT: Ugen, Kenneth E.  
; APPLICANT: Williams, William V.  
; TITLE OF INVENTION: Methods and Compositions for Diagnosing  
; TITLE OF INVENTION: and Treating Certain HIV Infected Patients  
; NUMBER OF SEQUENCES: 197  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson  
STREET: P.O. Box 457, 321 No. 55567441sttown Road  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/218,025A  
FILING DATE: 24-MAR-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/891,451  
FILING DATE: 29-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST33A  
TELEPHONE: (215) 540-9206  
TELEFAX: (215) 540-5818  
INFORMATION FOR SEQ ID NO: 105:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-218-025A-105

Query Match 27.6%; Score 26.5; DB 1; Length 15;  
Best Local Similarity 45.5%; Pred. No. 2e+02;  
Matches 5; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 4 CNDVTRLRLR 14  
| | | | |  
DB 1 CSFN-ISTSIR 10

RESULT 17  
US-09-224-785-22  
; Sequence 22; Application US/09224785A  
; Patent No. 6197526  
; GENERAL INFORMATION:  
; APPLICANT: Potter, M. Daniel  
; APPLICANT: Yu, Jinan  
; APPLICANT: Kelley, Brian D.  
; APPLICANT: Deetz, Jeffrey S.  
; APPLICANT: Booth, James E.  
; TITLE OF INVENTION: Binding Molecules for Human Factor VIII and Factor  
; FILE REFERENCE: Dyax-008 US sequence listing  
; CURRENT APPLICATION NUMBER: US/09/224,785A  
; CURRENT FILING DATE: 1999-01-04  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 22  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-224-785-22

Query Match 27.1%; Score 26; DB 4; Length 11;  
Best Local Similarity 60.0%; Pred. No. 1.7e+02;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 PERCH 5  
1:11:  
Db 7 FDHCH 11

RESULT 18  
US-09-224-785-25  
; Sequence 25, Application US/09224785A  
; Patent No. 6197526  
; GENERAL INFORMATION:  
; APPLICANT: Potter, M. Daniel  
; APPLICANT: Yu, Jinan  
; APPLICANT: Kelley, Brian D  
; APPLICANT: Deetz, Jeffrey S  
; TITLE OF INVENTION: Binding Molecules for Human Factor VIII and Factor  
; TITLE OF INVENTION: VIII-like Proteins  
; FILE REFERENCE: Dvax-008 US sequence listing  
; CURRENT APPLICATION NUMBER: US/09/224,785A  
; CURRENT FILING DATE: 1999-01-04  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 25  
; LENGTH: 11  
; TYPE: PRP  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-224-785-25

Query Match 27.1%; Score 26; DB 4; Length 11;  
Best Local Similarity 60.0%; Pred. No. 1.7e+02;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 PERCH 5  
1:11:  
Db 7 FDHCH 11

RESULT 19  
US-07-778-233B-33  
; Sequence 33, Application US/07778233B  
; Patent No. 5270170  
; GENERAL INFORMATION:  
; APPLICANT: Schatz, Peter J.  
; APPLICANT: Cull, Willard G.  
; APPLICANT: Miller, Jeff F.  
; TITLE OF INVENTION: Peptide Library and Screening Method  
; NUMBER OF SEQUENCES: 78  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: William M. Smith  
; STREET: One Market Plaza, Stewart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/778,233B  
; FILING DATE: 19911016  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 11509-50  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; CLONE: 7 4 1.1  
US-07-778-233B-33

Query Match 27.1%; Score 26; DB 1; Length 12;  
Best Local Similarity 55.6%; Pred. No. 1.9e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 9 VTPRLRENE 17  
1:1:1:1:  
Db 1 VTPRLRSNK 9

RESULT 20  
US-07-963-321-33  
; Sequence 33, Application US/07963321  
; Patent No. 538665  
; GENERAL INFORMATION:  
; APPLICANT: Schatz, Peter J.  
; APPLICANT: Cull, Willard G.  
; APPLICANT: Miller, Jeff F.  
; APPLICANT: Stemmer, Willem P.C.  
; TITLE OF INVENTION: Peptide Library and Screening Method  
; NUMBER OF SEQUENCES: 91  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: William M. Smith  
; STREET: One Market Plaza, Stewart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/963,321  
; FILING DATE: 19921015  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/778,223  
; FILING DATE: 16-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 11509-50-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; CLONE: 7 4 1.1  
US-07-963-321-33

Query Match 27.1%; Score 26; DB 1; Length 12;

Best Local Similarity 55.6%; Pred. No. 1.9e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 VTTRLRENE 17  
| | | | |  
Db 1 VTORVRNKK 9

## RESULT 21

US-08-290-641-33  
; Sequence 33, Application US/08290641  
; Patent No. 5498530

## ; GENERAL INFORMATION:

APPLICANT: Schatz, Peter J.  
APPLICANT: Cull, Millard G.  
APPLICANT: Miller, Jeff F.  
APPLICANT: Stemmer, William P.C.  
TITLE OF INVENTION: Peptide Library and Screening Method  
NUMBER OF SEQUENCES: 91  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: William M. Smith  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/290,641  
FILING DATE: 15-AUG-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,321  
FILING DATE: 15-OCT-1992  
APPLICATION NUMBER: US 07/778,223  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11509-50-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
CLONE: 7 4 1.1  
US-08-290-641-33

## Query Match

Best Local Similarity 27.1%; Score 26; DB 1; Length 12;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 VTTRLRENE 17  
| | | | |  
Db 1 VTORVRNKK 9

## RESULT 22

US-08-548-540-33  
; Sequence 33, Application US/08548540  
; Patent No. 5733731  
; GENERAL INFORMATION:

APPLICANT: Schatz, Peter J.  
APPLICANT: Cull, Millard G.  
APPLICANT: Miller, Jeff F.  
APPLICANT: Stemmer, William P.C.  
TITLE OF INVENTION: Peptide Library and Screening Method  
NUMBER OF SEQUENCES: 162  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: William M. Smith  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/548,540  
FILING DATE: 26-OCT-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/290,641  
FILING DATE: 15-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,321  
FILING DATE: 15-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 16528J-001240US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
CLONE: 7 4 1.1  
US-08-548-540-33

## Query Match

Best Local Similarity 27.1%; Score 26; DB 1; Length 12;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 VTTRLRENE 17  
| | | | |  
Db 1 VTORVRNKK 9

## RESULT 23

PCT-US96-09809-33  
; Sequence 33, Application PC/TUS9609809  
; GENERAL INFORMATION:

APPLICANT: Schatz, Peter J.  
APPLICANT: Cull, Millard G.  
APPLICANT: Miller, Jeff F.  
APPLICANT: Stemmer, William P.C.  
TITLE OF INVENTION: Peptide Library and Screening Method  
NUMBER OF SEQUENCES: 162  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: William M. Smith  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California

COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/09809  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/548,540  
FILING DATE: 26-OCT-1995  
APPLICATION NUMBER: US 08/290,641  
FILING DATE: 15-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,321  
FILING DATE: 15-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 16528J-001240US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
CLONE: 741.1  
PCT-US96-09809-33

Query Match 27.1%; Score 26; DB 5; Length 12;  
Best Local Similarity 55.6%; Pred. No. 1.9e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 VTTRLRNE 17  
| | | | |  
DB 1 VTQVRVSNK 9

RESULT 24  
US-08-413-708B-5  
Sequence 5, Application US/08413708B  
Patent No. 5972883  
GENERAL INFORMATION:  
APPLICANT: GOZES, Iliana  
APPLICANT: FRIDKIN, Matlyahu  
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING  
TITLE OF INVENTION: VIP, OR ANALOGUES DERIVATIVES AND FRAGMENTS THEREOF  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEWMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/413,708B  
FILING DATE: 30-MAR-1995  
CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/207,671  
FILING DATE: 09-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IL 105061  
FILING DATE: 16-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BROWDY, Roger L.  
REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: GOZES-3A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-413-708B-5

Query Match 27.1%; Score 26; DB 2; Length 14;  
Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 FNDVTRLR 14  
| | | | |  
DB 6 FTDNVTRLR 14

RESULT 25  
US-09-192-048-24  
Sequence 24, Application US/09192048  
Patent No. 6063592  
GENERAL INFORMATION:  
APPLICANT: Lee, Soohae  
TITLE OF INVENTION: Kell Protein Proteolytic Activity  
FILE REFERENCE: Sequence 1-27  
Patent No. 6063592  
CURRENT APPLICATION NUMBER: US/09/192,048  
CURRENT FILING DATE: 1998-11-13  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 24  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-192-048-24

Query Match 27.1%; Score 26; DB 3; Length 14;  
Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 FNDVTRLR 14  
| | | | |  
DB 6 FTDNVTRLR 14

Search completed: April 2, 2002, 09:24:10  
Job time: 79 sec

Tue Apr 2 09:51:48 2002

us-09-020-393b-3\_copy\_42\_58.closed.ra

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:09:12 ; Search time 19.6 Seconds  
(without alignments)  
104.934 Million cell updates/sec

Title: US-09-020-393b-18

Perfect score: 145

Sequence: 1 KCLGYHLVDVSLAFSEISVGAERFKDDC 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 5455

Minimum DB seq length: 0

Maximum DB seq length: 27

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	21.4	22	2	T25653
2	28	19.3	17	2	A60889
3	28	19.3	22	2	A36389
4	28	19.3	27	2	A41006
5	27	18.6	22	2	JP0066
6	27	18.6	25	2	S07574
7	26.5	18.3	20	2	S72501
8	26	17.9	15	2	S08301
9	26	17.9	18	2	A61382
10	26	17.9	21	2	I65270
11	26	17.9	24	1	B32252
12	26	17.9	24	2	T46622
13	26	17.9	25	2	A48810
14	25	17.2	20	2	A60295
15	25	17.2	27	2	T37261
16	24	16.6	17	2	S71864
17	24	16.6	17	2	A36727
18	24	16.6	21	2	S38739
19	24	16.6	22	2	C46285
20	24	16.6	22	2	PC2134
21	24	16.6	23	2	S38738
22	23	15.9	14	2	PS0371
23	23	15.9	14	2	B61597
24	23	15.9	15	2	PA0062
25	23	15.9	15	2	D56385
26	23	15.9	15	2	A47628
27	23	15.9	16	2	B45895
28	23	15.9	17	2	PH1357
29	23	15.9	18	2	S36121

30	23	15.9	19	2	A38386	Ca2+-transporting
31	23	15.9	22	2	A35418	brain natriuretic
32	23	15.9	22	2	JT0581	natriuretic peptid
33	23	15.9	24	1	S58242	pyrroloquinoline q
34	23	15.9	24	2	S77982	cytochrome-c oxida
35	23	15.9	25	2	S39360	CDK inhibitor - mo
36	23	15.9	26	2	S59906	gamma-glutamyl tra
37	23	15.9	26	2	D53507	anexin V - rat (f
38	23	15.9	27	2	JC1081	brain natriuretic
39	22.5	15.5	26	4	T07533	probable photosyst
40	22	15.2	11	2	A35594	buccalin - Callor
41	22	15.2	15	2	S66215	cartilage oligomer
42	22	15.2	20	2	S23981	outer layer protei
43	22	15.2	20	2	A61526	major milk gland p
44	22	15.2	21	2	F64121	hypothetical prote
45	22	15.2	23	1	S20453	pyrroloquinoline q

#### ALIGNMENTS

RESULT 1  
T25653  
hypothetical protein C47C12.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T25653  
R:Connell, M.  
submitted to the EMBL Data Library, August 1996  
A:Description: The sequence of C. elegans cosmid C47C12.  
A:Reference number: Z20062  
A:Accession: T25653  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-22 <CON>  
A:Cross-references: EMBL:U67951; PIDN:AB07573.1; GSPDB:GN00028; CESP:C47C12.5  
A:Experimental source: strain Bristol N2; clone C47C12  
C:Genetics:  
A:Gene: CESP:C47C12.5  
A:Map position: X  
A:Introns: 21/1

Query Match 21.4%; Score 31; DB 2; Length 22;  
Best Local Similarity 50.0%; Pred. No. 2.4e+02;  
Matches 9; Conservative 2; Mismatches 5; Indels 2; Gaps 1;  
OY 6 HLDV--SLAFSEISVGAE 21  
| : | | | | : |  
Db 4 HSDIGNSLOFSRISLKNE 21  
RESULT 2  
A60889  
olfactory glycoprotein RB-8 - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: A60889  
R:Schwab, J.E.; Gottlieb, D.I.  
J. Neurosci. 8, 3470-3480, 1988  
A:Title: Purification and characterization of an antigen that is spatially segregated  
A:Reference number: A60889; MUID:89010968  
A:Accession: A60889  
A:Molecule type: protein  
A:Residues: 1-17 <SCH>  
A:Note: this protein was purified from whole brain  
C:Comment: The monoclonal antibody RB-8 binds this integral membrane glycoprotein on  
C:Keywords: glycoprotein; membrane protein; olfaction

Query Match 19.3%; Score 28; DB 2; Length 17;  
Best Local Similarity 54.5%; Pred. No. 5.5e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 9 VSLAFSEISVG 19  
 :||: |||||  
 Db 6 ISLSKVELSVG 16

RESULT 3  
 A36399  
 C-type natriuretic peptide - frog  
 C:Species: Ranidae gen. sp. (frog)  
 C>Date: 01-Feb-1991 #sequence\_revision 01-Feb-1991 #text\_change 09-Dec-1994  
 C:Accession: A36399  
 R:Yoshihara, A.; Kozawa, H.; Minamino, N.; Kangawa, K.; Matsuo, H.  
 Biochem. Biophys. Res. Commun. 173, 591-598, 1990  
 A:Title: Isolation and sequence determination of frog C-type natriuretic peptide.  
 A:Reference number: A36399; MUID:91083642  
 A:Accession: A36399  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-22 <YOS>  
 C:Superfamily: natriuretic peptide A precursor

Query Match 19.3%; Score 28; DB 2; Length 22;  
 Best Local Similarity 53.8%; Pred. No. 7.3e+02;  
 Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 OY 2 CLGYHDLVSLAFS 14  
 :||: |||||  
 Db 6 CFCVGLDKRIGAFS 18

RESULT 4  
 A41006  
 LDL receptor - chicken (fragment)  
 C:Species: Gallus gallus (chicken)  
 C>Date: 12-Jun-1992 #sequence\_revision 12-Jun-1992 #text\_change 23-Mar-1993  
 C:Accession: A41006  
 R:Barber, D.L.; Sanders, E.J.; Aebersold, R.; Schneider, W.J.  
 J. Biol. Chem. 266, 18761-18770, 1991  
 A:Title: The receptor for yolk lipoprotein deposition in the chicken oocyte.  
 A:Reference number: A41006; MUID:92011658  
 A:Accession: A41006  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-27 <BAK>

Query Match 19.3%; Score 28; DB 2; Length 27;  
 Best Local Similarity 31.8%; Pred. No. 9.1e+02;  
 Matches 7; Conservative 2; Mismatches 13; Indels 0; Gaps 0;  
 OY 5 YHLDVSLAFSEISVGAEFNKDD 26  
 :||: |||||  
 Db 3 YWDTSSAIFSAISIDTSMNFDN 24

RESULT 5  
 JP0066  
 ribosomal protein L30 - Nocardia asteroides (fragment)  
 C:Species: Nocardia asteroides  
 C>Date: 10-Mar-1994 #sequence\_revision 28-Oct-1994 #text\_change 28-Oct-1994  
 C:Accession: JP0066  
 R:Ochi, K.  
 submitted to JPIPD, February 1994  
 A:Description: Phylogenetic diversity in the genus Bacillus and comparative ribosomal p  
 A:Reference number: JP0066  
 A:Accession: JP0066  
 A:Molecule type: protein  
 A:Residues: 1-22 <OCH>  
 C:Keywords: protein biosynthesis; ribosome

Query Match 18.6%; Score 27; DB 2; Length 22;  
 Best Local Similarity 26.3%; Pred. No. 1e+03;  
 Matches 5; Conservative 7; Mismatches 7; Indels 0; Gaps 0;  
 OY 8 DVSLAFSEISVGAEFNKDD 26  
 :||: |||||  
 Db 2 DLKVTQIKSTIGAKANKDD 20

RESULT 6  
 S07574  
 malate dehydrogenase (EC 1.1.1.37) - Phenyllobacterium immobile (fragment)  
 C:Species: Phenyllobacterium immobile  
 C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 02-Aug-1994  
 C:Accession: S07574  
 R:Rommel, T.O.; Hurd, H.K.; Speth, A.R.; Lingens, F.  
 Biol. Chem. Hoppe-Seyler 370, 763-768, 1989  
 A:Title: Purification and N-terminal amino-acid sequences of bacterial malate dehydr  
 A:Reference number: S04956; MUID:89374824  
 A:Accession: S07574  
 A:Molecule type: protein  
 A:Residues: 1-25 <ROM>  
 C:Superfamily: L-lactate dehydrogenase  
 C:Keywords: oxidoreductase; tricarboxylic acid cycle

Query Match 18.6%; Score 27; DB 2; Length 25;  
 Best Local Similarity 80.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 3 LGYHL 7  
 :|||  
 Db 16 IGYHL 20

RESULT 7  
 S72501  
 protein kinase C inhibitor - human (fragment)  
 N:Alternate names: histidine triad nucleotide-binding protein  
 C:Species: Homo sapiens (man)  
 C>Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999  
 C:Accession: S72501; S62623  
 R:Maines, M.D.; Trakshel, G.M.  
 Arch. Biochem. Biophys. 300, 320-326, 1993  
 A:Title: Purification and characterization of human biliverdin reductase.  
 A:Reference number: S29736; MUID:9314333  
 A:Accession: S72501  
 A:Molecule type: protein  
 A:Residues: 1-20 <MAI>  
 A:Note: this protein was identified as biliverdin reductase; the identification is qu  
 R:Maines, M.D.; Polevoda, B.V.; Huang, T.J.; McCoubrey Jr., W.K.  
 Eur. J. Biochem. 235, 372-381, 1996  
 A:Title: Human biliverdin IX-alpha reductase is a zinc-metalloprotein. Characterizati  
 A:Reference number: S62622; MUID:96202961  
 A:Accession: S62623  
 A:Molecule type: protein  
 A:Residues: 1-20 <MAW>  
 C:Superfamily: protein kinase C inhibitor; histidine triad homology  
 C:Keywords: homodimer; protein kinase inhibitor; zinc

Query Match 18.3%; Score 26.5; DB 2; Length 20;  
 Best Local Similarity 55.6%; Pred. No. 1.1e+03;  
 Matches 5; Conservative 2; Mismatches 1; Indels 1; Gaps 1;  
 OY 2 CLGYHDLVS 10  
 :||: |||||  
 Db 1 CLAFH-DIS 8

RESULT 8  
 S08301  
 epidermal growth factor, high molecular weight - rat (fragment)



C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 30-Sep-1993  
C:Accession: S08301  
R:Nexo, E.; Jorgensen, P.E.; Thim, L.; Roepstorff, P.  
Biochim. Biophys. Acta 1037, 388-393, 1990  
A:Title: Purification and characterization of a low and a high molecular weight form of  
A:Reference number: S08288; MUID:90181442  
A:Accession: S08301  
A:Molecule type: protein  
A:Residues: 1-15 <NEX>  
C:Keywords: growth factor

Query Match 17.9%; Score 26; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 KDCC 27  
| | | |  
Db 3 KDCC 6

RESULT 9  
A61392  
brain-associated small cell lung cancer antigen - human (fragment)  
N:Alternate names: BASCA  
C:Species: Homo sapiens (man)  
C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 09-Sep-1994  
C:Accession: A61392 - sequence\_revision 09-Sep-1994 #text\_change 09-Sep-1994  
R:Umezawa, Y.; Kuge, S.; Kikyo, N.; Shirai, T.; Watanabe, J.; Fujiwara, M.; Okabe, T.  
Jpn. J. Clin. Oncol. 21, 251-255, 1991  
A:Title: Identity of brain-associated small cell lung cancer antigen and the CD56 (NKH-1  
A:Reference number: A61392; MUID:92046737  
A:Accession: A61392  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-18 <DME>

Query Match 17.9%; Score 26; DB 2; Length 18;  
Best Local Similarity 46.2%; Pred. No. 1.2e+03;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 7 LDVSLAFSEISVG 19  
| | | | | | | | | |  
Db 3 VDIVPSGGEISVG 15

RESULT 10  
I65270  
collagen alpha 1(I) chain - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 13-Aug-1999  
C:Accession: I65270  
R:Genovesse, C.; Rowe, D.; Kream, B.  
Biochemistry 23, 6210-6216, 1984  
A:Title: Construction of DNA sequences complementary to rat alpha-1 and alpha-2 collagen  
A:Reference number: I52392; MUID:85122694  
A:Accession: I65270  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-21 <RES>  
A:Cross-references: GB:M12200; NID:g203191; PIDN:AAA40835.1; PID:g203195  
C:Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology

Query Match 17.9%; Score 26; DB 2; Length 21;  
Best Local Similarity 33.3%; Pred. No. 1.4e+03;  
Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 5 YHLDVSLAFSEISVG 19  
| | | | | | | | | |  
Db 7 YHCKNSIAYLDEMG 21

RESULT 11  
B32252  
pyrroloquinoline quinone precursor - Acinetobacter calcoaceticus  
N:Alternate names: pqg gene IV protein  
C:Species: Acinetobacter calcoaceticus  
C:Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 18-Feb-2000  
C:Accession: B32252  
R:Goosen, N.; Horsman, H.P.A.; Huinen, R.G.M.; van de Putte, P.  
J. Bacteriol. 171, 447-455, 1989  
A:Title: Acinetobacter calcoaceticus genes involved in biosynthesis of the coenzyme P  
A:Reference number: A32252; MUID:89123056  
A:Accession: B32252  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-24 <GOO>  
A:Cross-references: GB:X06452; NID:g38740; PID:g130796  
A:Note: this ORF is not annotated in Genbank entry ACPQ05YN, release 116.0  
R:Goosen, N.; Huinen, R.G.; van de Putte, P.  
J. Bacteriol. 174, 1426-1427, 1992  
A:Title: A 24-amino-acid polypeptide is essential for the biosynthesis of the coenzyme  
A:Reference number: A59183; MUID:92138642  
A:Contents: annotation  
C:Genetics:  
A:Gene: pqgA  
C:Superfamily: pyrroloquinoline quinone precursor pqgA  
C:Keywords: quinoprotein  
F:16/20/Product: pyrroloquinoline quinone #status predicted <MAN>  
F:16-20/Cross-Link: pyrroloquinoline quinone (Glu, Tyr) #status predicted

Query Match 17.9%; Score 26; DB 1; Length 24;  
Best Local Similarity 40.0%; Pred. No. 1.6e+03;  
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 12 AFSEISVGAE 21  
| | | | | | | | | |  
Db 7 AFTDLRIGFE 16

RESULT 12  
T46622  
hypothetical protein cl - lobliolly pine  
C:Species: Pinus taeda (loblolly pine)  
C:Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 18-Feb-2000  
C:Accession: T46622  
R:Chang, S.; Puryear, J.; Funkhouser, E.A.; Newton, R.J.; Cairney, J.  
submitted to the EMBL Data Library, July 1995  
A:Description: Cloning of a chitinase homolog which lacks chitin binding sites and is  
A:Reference number: 223105  
A:Accession: T46622  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-24 <CHA>  
A:Cross-references: EMBL:U01309; NID:g974285; PID:g974287  
A:Experimental source: strain s6PT2xs6PT3; 8 month seedlings

Query Match 17.9%; Score 26; DB 2; Length 24;  
Best Local Similarity 41.7%; Pred. No. 1.6e+03;  
Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 2 CLGYHLDVSLAF 13  
| | | | | | | | | |  
Db 11 CGFQPPDCLICF 22

RESULT 13  
A48810  
fibrinogen B beta subunit - African clawed frog (fragment)  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000

C:Accession: A48810  
 R:Roberts, L.R.; Nichols, L.A.; Holland, L.J.  
 Biochemistry 32, 11627-11637, 1993  
 A:Title: Transcriptional regulation of the Xenopus laevis B beta fibrinogen subunit gene  
 A:Reference number: A48810; MUID:94032285  
 A:Accession: A48810  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-25 <ROB>  
 A:Cross-references: GB:U05035; GB:S66373; NID:9450950; PIDN:AAA60463.1; PID:9450951  
 A>Note: sequence extracted from NCBI backbone (NCBIN:138880, NCBI:138881)

Query Match 17.9%; Score 26; DB 2; Length 25;  
 Best Local Similarity 23.5%; Pred. No. 1.7e+03;  
 Matches 4; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 10 SLAFSEISVGAEPNKDD 26  
 :| | : : : : :|  
 Db 8 ALCVSAWCSVDYEDD 24

RESULT 14  
 A60295  
 apolipophorin III - house cricket (fragment)  
 C:Species: Acheta domestica (house cricket)  
 C:Date: 03-Feb-1993 #sequence\_revision 03-Feb-1993 #text\_change 18-Jun-1993  
 A:Accession: A60295  
 R:Strobel, L.M.; Kanost, M.R.; Ziegler, R.; Wells, M.A.  
 Insect Biochem. 20, 859-863, 1990  
 A:Title: Adipokinetic hormone causes formation of a low density lipophorin in the house  
 A:Reference number: A60295  
 A:Accession: A60295  
 A:Molecule type: protein  
 A:Residues: 1-20 <STR>  
 C:Comment: This protein, a small, water-soluble apolipophorin, is thought to increase the  
 residual pathway for flight-related lipid transport.  
 C:Keywords: hemolymph; lipid transport

Query Match 17.2%; Score 25; DB 2; Length 20;  
 Best Local Similarity 57.1%; Pred. No. 1.9e+03;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 17 SVGAEPN 23  
 :| | : : : : :|  
 Db 4 TTGADEN 10

RESULT 15  
 I37261  
 complement receptor - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 21-Jul-2000  
 A:Accession: I37261  
 R:Bitenbach, M.; Tong, X.; Bradbury, L.E.; Tedder, T.F.; Kieff, E.  
 J. Exp. Med. 176, 1405-1414, 1992  
 A:Title: Characterization of an Epstein-Barr virus receptor on human epithelial cells.  
 A:Reference number: I37261; MUID:93018869  
 A:Accession: I37261  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-27 <RES>  
 A:Cross-references: EMBL:X68990; NID:93928195; PIDN:CAA48779.1; PID:9395334  
 C:Genetics:  
 A:Gene: GDB:CR2  
 A:Cross-references: GDB:119802; OMIM:120650  
 A:Map position: Iq32-Iq32

Query Match 17.2%; Score 25; DB 2; Length 27;  
 Best Local Similarity 33.3%; Pred. No. 2.7e+03;  
 Matches 4; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 CLGYHLYSLAF 13  
 | | : | :|  
 Db 6 CAGHYLRRLSM 17

RESULT 16  
 S71864  
 glutathione transferase (EC 2.5.1.18) class alpha 6a - pig (fragment)  
 N:Alternate names: glutathione S-transferase class alpha 6a  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 19-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 07-May-1999  
 A:Accession: S71864  
 R:Round, P.; Anglade, P.; Debrauer, L.; Tulliez, J.  
 Biochem. J. 317, 879-884, 1996  
 A:Title: Characterization of pig liver glutathione S-transferases using HPLC-electros  
 A:Reference number: S71864; MUID:96332484  
 A:Accession: S71864  
 A:Molecule type: protein  
 A:Residues: 1-17 <ROU>  
 A:Experimental source: liver; cytosolic  
 C:Comment: At least five species-independent classes of cytosolic glutathione transfer  
 s mitochondrial form are known.  
 C:Function: dimer  
 A:Description: catalyzes the nucleophilic conjugation of intracellular glutathione to  
 A:Pathway: detoxification; xenobiotics metabolism  
 A>Note: Increased hydrophilicity of GSH-conjugates facilitates their further metabol  
 es of damage  
 C:Superfamily: glutathione transferase  
 C:Keywords: dimer; transferase

Query Match 16.6%; Score 24; DB 2; Length 17;  
 Best Local Similarity 35.7%; Pred. No. 2.3e+03;  
 Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 8 DVSLAFSEISVGA 21  
 | | | : :|  
 Db 1 DGSLLFQVDMYTE 14

RESULT 17  
 A36727  
 cytochrome c551 - Methylomonas sp. (fragment)  
 C:Species: Methylomonas sp.  
 C:Date: 19-Apr-1991 #sequence\_revision 19-Apr-1991 #text\_change 18-Jun-1993  
 A:Accession: A36727  
 R:DiSpirito, A.A.; Lipscomb, J.D.; Lidstrom, M.E.  
 J. Bacteriol. 172, 5360-5367, 1990  
 A:Title: Soluble cytochromes from the marine methanotroph Methylomonas sp. strain A4.  
 A:Reference number: A36727; MUID:90366596  
 A:Accession: A36727  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-17 <DIS>

Query Match 16.6%; Score 24; DB 2; Length 17;  
 Best Local Similarity 60.0%; Pred. No. 2.3e+03;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CLGYH 6  
 | | :|  
 Db 13 CSGFH 17

RESULT 18  
 S38739  
 lipid transfer protein S2 - spinach (fragment)  
 C:Species: Spinacia oleracea (spinach)  
 C:Date: 08-Jun-1994 #sequence\_revision 19-Jan-1996 #text\_change 03-May-1996  
 A:Accession: S38739

R;Segura, A.; Moreno, M.; Garcia-Olmedo, F.  
FEBS Lett. 332, 243-246, 1993  
A:Title: Purification and antipathogenic activity of lipid transfer proteins (LTRs) from  
A:Reference number: S38736; MUID:94009709  
A:Accession: S38739  
A:Molecule type: protein  
A:Residues: 1-21 <SEG>  
C:Superfamily: phospholipid transfer protein

Query Match 16.6%; Score 24; DB 2; Length 21;  
Best Local Similarity 75.0%; Pred. No. 2.9e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CIGY 5  
|:|:  
Db 14 CIGY 17

RESULT 19  
C46285  
formaldehyde dehydrogenase (glutathione) (EC 1.2.1.1) class III low activity form - Atla  
C:Species: Gadus morhua (Atlantic cod)  
C:Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Mar-1996  
C:Accession: C46285  
R:Danielsson, O.; Jornvall, H.  
Proc. Natl. Acad. Sci. U.S.A. 89, 9247-9251, 1992  
A:Title: "Enzymogenesis": classical liver alcohol dehydrogenase origin from the glutath  
A:Reference number: A46285; MUID:93028441  
A:Accession: C46285  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-22 <DAN>  
A:Note: sequence extracted from NCBI backbone (NCBIP:116269)  
C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology  
C:Keywords: NAD; oxidoreductase

Query Match 16.6%; Score 24; DB 2; Length 22;  
Best Local Similarity 50.0%; Pred. No. 3.1e+03;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 20 AERKDC 27  
|:|:  
Db 11 AERKDC 18

RESULT 20  
PC2134  
maltose transport protein - yeast (Saccharomyces cerevisiae) (fragment)  
C:Species: Saccharomyces cerevisiae  
C:Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #text\_change 07-May-1999  
C:Accession: PC2134  
R:Van den Broek, P.J.A.; Van Leeuwen, C.C.M.; Weusthuis, R.A.; Postma, E.; Van Dijken, J.  
Biochem. Biophys. Res. Commun. 200, 45-51, 1994  
A:Title: Identification of the maltose transport protein of Saccharomyces cerevisiae.  
A:Reference number: PC2133; MUID:94220124  
A:Accession: PC2134  
A:Molecule type: protein  
A:Residues: 1-22 <VAN>  
A:Experimental source: Plasma membrane  
C:Superfamily: maltose transport protein MAL61  
C:Keywords: membrane protein

Query Match 16.6%; Score 24; DB 2; Length 22;  
Best Local Similarity 35.7%; Pred. No. 3.1e+03;  
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 13 FSEISVGAERNKD 26  
|:|:  
Db 7 FAAAAAAEINVKD 20

RESULT 21  
S38738  
lipid transfer protein S1 - spinach (fragment)  
C:Species: Spinacia oleracea (spinach) 19-Jan-1996 #text\_change 03-May-1996  
C:Accession: S38738  
R;Segura, A.; Moreno, M.; Garcia-Olmedo, F.  
FEBS Lett. 332, 243-246, 1993  
A:Title: Purification and antipathogenic activity of lipid transfer proteins (LTRs) f  
A:Reference number: S38736; MUID:94009709  
A:Accession: S38738  
A:Molecule type: protein  
A:Residues: 1-23 <SEG>  
C:Superfamily: phospholipid transfer protein

Query Match 16.6%; Score 24; DB 2; Length 23;  
Best Local Similarity 75.0%; Pred. No. 3.2e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CIGY 5  
|:|:  
Db 14 CIGY 17

RESULT 22  
PS0371  
hypothetical protein (psac region) - Synechococcus sp. (fragment)  
C:Species: Synechococcus sp.  
C:Date: 17-Aug-1992 #sequence\_revision 17-Aug-1992 #text\_change 08-Oct-1999  
C:Accession: PS0371  
R:Riele, E.; Stirewalt, V.L.; Gasparich, G.E.; Bryant, D.A.  
Gene 112, 123-128, 1992  
A:Title: The psac genes of Synechococcus sp. PCC7002 and Cyanophora paradoxa: cloning  
A:Reference number: JS0694; MUID:92201692  
A:Accession: PS0371  
A:Molecule type: DNA  
A:Residues: 1-14 <RH1>  
A:Cross-references: GB:86238; NID:9154574; PIDN:AAA27351.1; PID:9552030

Query Match 15.9%; Score 23; DB 2; Length 14;  
Best Local Similarity 57.1%; Pred. No. 2.7e+03;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 GYHDSV 10  
|:|:  
Db 3 GFKDVT 9

RESULT 23  
B61597  
cytochrome P450 AL-2 - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 06-Jan-1995  
C:Accession: B61597  
R:Shimeno, H.; Toda, A.; Ogata, S.; Nagamatsu, A.  
Drug Metab. Dispos. 19, 291-297, 1991  
A:Title: Purification and aminopyrine monooxygenase activity of liver microsomal cyto  
A:Reference number: A61597; MUID:91292910  
A:Accession: B61597  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-14 <SH1>

Query Match 15.9%; Score 23; DB 2; Length 14;  
Best Local Similarity 38.5%; Pred. No. 2.7e+03;  
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 7 LDVSLAFSEISVG 19  
|:|:  
Db 7 LDVSLAFSEISVG 19

Db 1 MDLLSLSPFLNG 13

# RESULT 24

PA0062  
 fumarate hydratase (EC 4.2.1.2) - fungus (Fusarium sporotrichioides) (fragment)  
 C/Species: Fusarium sporotrichioides  
 C/Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Mar-2001  
 C/Accession: PA0062  
 R/Chow, L.P.; Fukaya, N.; Suglira, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.  
 submitted to JIPID, October 1994  
 A/Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporotrichioides  
 A/Reference number: PA0051  
 A/Accession: PA0062  
 A/Molecule type: protein  
 A/Residues: 1-15 <CHO>  
 C/Keywords: carbon-oxygen lyase; hydro-lyase

## Query Match

Best Local Similarity 15.9%; Score 23; DB 2; Length 15;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 12 AFSEISV 18

Db 9 AFGEIOV 15

# RESULT 25

D56385  
 nitrophorin 4 - Rhodnius prolixus (fragment)  
 N/Alternate names: nitric oxide-carry protein NP4  
 C/Species: Rhodnius prolixus  
 C/Date: 21-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 17-Mar-1999  
 C/Accession: D56385  
 R/Champagne, D.E.; Nussenzweig, R.H.; Ribeiro, J.M.C.  
 J. Biol. Chem. 270, 8691-8695, 1995  
 A/Title: Purification, partial characterization, and cloning of nitric oxide-carrying heme  
 A/Reference number: A56385; MUID:95238361  
 A/Accession: D56385  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 1-15 <CHA>  
 C/Keywords: heme; salivary gland; vasodilator

## Query Match

Best Local Similarity 15.9%; Score 23; DB 2; Length 15;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 FNKD 25

Db 12 FNKD 15

Search completed: January 6, 2002, 09:13:53  
 Job time: 281 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:13:07 ; Search time 12.95 Seconds  
(without alignments)  
76.444 Million cell updates/sec

Title: US-09-020-393b-18

Perfect score: 145  
Sequence: 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27

Scoring table: BIOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 1552

Minimum DB seq length: 0  
Maximum DB seq length: 27

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	18.6	25	1 MDH_PHEIM	P19980 phenylolact
2	26	17.9	24	1 POQA_ACICA	P27532 acinetobact
3	25	17.2	20	1 TPX_CLOPA	P81361 clostridium
4	24	16.6	15	1 RML2_YEAST	P36522 saccharomyc
5	23	15.9	12	1 XYLA_STRVN	P14405 streptomyc
6	23	15.9	13	1 BP37_LEUMA	P81754 leucophaea
7	23	15.9	18	1 AGI_EUPCH	P33888 euphorbia c
8	23	15.9	22	1 ANFC_CHICK	P21805 gallus gall
9	23	15.9	23	1 FMK7_PSEAE	P53391 pseudomonas
10	23	15.9	24	1 COXC_THUOB	P80973 thunnus obe
11	23	15.9	24	1 POQA_PSEFL	P55171 pseudomonas
12	23	15.9	27	1 KT39_PICKL	P80326 picchia kluy
13	22	15.2	20	1 VMO2_CHICK	O99849 gallus gall
14	22	15.2	21	1 YD90_HAEIN	P45194 haemophilus
15	22	15.2	23	1 POQA_KLEPN	P27503 klebsiella
16	22	15.2	25	1 PETM_SKECO	O96810 skeletonema
17	21.5	14.8	27	1 CX7A_CONTE	P24160 conus texti
18	21	14.5	10	1 TK14_LOCOM	P30250 locusta mig
19	21	14.5	19	1 FIBR_SHEEP	P14451 ovis aries
20	21	14.5	22	1 VI01_VACCV	P16714 vaccinia vi
21	21	14.5	22	1 RL10_KLEPN	P41190 klebsiella
22	21	14.5	23	1 RL10_ENTCL	O47608 enterobacte
23	21	14.5	24	1 CAMT_PINPS	P81081 pinus pinas
24	21	14.5	27	1 ANF_ANGJA	P18144 anguilla ja
25	21	14.5	27	1 CD55_SHEEP	P58020 ovis aries
26	21	14.5	27	1 CXK7_CONPU	P56633 conus purpu
27	20.5	14.1	10	1 YMCD_EMENT	P03884 emeritocella
28	20	13.8	17	1 SP34_DICMU	P81545 dicyostella
29	20	13.8	17	1 NU4M_TRIRU	O36834 trichophyto
30	20	13.8	18	1 AGI_EUPMA	P33889 euphorbia m
31	20	13.8	21	1 UP04_CAEEL	P55955 caenorhabdi
32	20	13.8	24	1 ATPB_MICLU	P80266 micrococcus
33	20	13.8	24	1 PL13_PLETR	P36987 plectreury

34	20	13.8	25	1 GBX1_MOUSE	P82976 mus musculu
35	20	13.8	27	1 DBH1_BIFLO	P17615 bifidobacte
36	19.5	13.4	21	1 BTX_ATRBI	P80163 atractaspis
37	19.5	13.4	27	1 RL22_HALCU	P05973 halobacteri
38	19	13.1	10	1 TRP8_LEUMA	P81740 leucophaea
39	19	13.1	12	1 V25K_WSSV	P82004 white spot
40	19	13.1	14	1 MAST_POLJA	P01517 polistes ja
41	19	13.1	15	1 PH2_PERAM	P82695 periplaneta
42	19	13.1	16	1 LE06_BIOGL	P80745 biomphalari
43	19	13.1	16	1 RIPK_TRIKI	P16093 trichosan
44	19	13.1	17	1 PH3_PERAM	P82696 periplaneta
45	19	13.1	17	1 PH4_PERAM	P82697 periplaneta

## ALIGNMENTS

RESULT	ID	MDH_PHEIM	STANDARD	PRT	25 AA.
AC	P19980				
DT	01-FEB-1991 (Rel. 17, Created)				
DT	01-FEB-1991 (Rel. 17, Last sequence update)				
DT	30-MAY-2000 (Rel. 39, Last annotation update)				
DE	MALATE DEHYDROGENASE (EC 1.1.1.37) (FRAGMENT)				
GN	MDH.				
OS	Phenyllobacterium immobile.				
OC	Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;				
OC	Phenyllobacterium.				
OX	NCBI_TaxID=21;				
RN	[1]				
RP	SEQUENCE.				
RC	STRAIN=E;				
RX	MEDLINE=89374824; PubMed=2775496;				
RA	Rommel T.O., Hund H.-K., Specht A.R., Lingens F.;				
RT	"Purification and N-terminal amino-acid sequences of bacterial malate				
RT	dehydrogenases from six actinomycetales strains and from				
RT	phenyllobacterium immobile, strain E.";				
RL	Biol. Chem. Hoppe-Seyler 370:763-768(1989).				
CC	-1- CATALYTIC ACTIVITY: L-MALATE + NAD(+) -> OXALOACETATE + NADH.				
CC	-1- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.				
CC	PIR: S07574; S07574.				
DR	HSSP: P10584; IBDM.				
DR	InterPro: IPR001252; MDH_actsite.				
DR	InterPro: IPR001236; ldh.				
DR	Pfam: PF00056; ldh; 1.				
DR	PROSITE: PS00068; MDH; PARTIAL.				
KW	Oxidoreductase; Tricarboxylic acid cycle; NAD.				
FT	NON_TER				
FT	SEQUENCE 25 AA; 2626 MW; C8D81E008825845C CRC64;				
QY	3 LGYHL 7				
QY	16 IGYHL 20				
DB	16 IGYHL 20				
RESULT	2				
ID	POQA_ACICA	STANDARD;	PRT;	24 AA.	
AC	P27532;				
DT	01-AUG-1992 (Rel. 23, Created)				
DT	01-AUG-1992 (Rel. 23, Last sequence update)				
DT	01-OCT-1996 (Rel. 34, Last annotation update)				
DE	COENZYME PQQ SYNTHESIS PROTEIN A (COENZYME PQQ SYNTHESIS PROTEIN IV).				
GN	POQA OR PQQ IV.				
OS	Acinetobacter calcoaceticus.				
OC	Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;				
OC	Acinetobacter.				

OX NCBI\_TaxID=471;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LMD 79.41;  
 RX MEDLINE=89123056; PubMed=2536663;  
 RA Goosen N., Horsman H.P.A., Huinen R.G.M., van de Putte P.;  
 RT "Acinetobacter calcoaceticus genes involved in biosynthesis of the  
 coenzyme pyruvate-quinoline-quinone: nucleotide sequence and  
 expression in *Escherichia coli* K-12.";  
 RL J. Bacteriol. 171:447-455(1989).  
 CC -1- FUNCTION: REQUIRED FOR COENZYME PYRUVATE-QUINOLINE-QUINONE (PQQ)  
 BIOSYNTHESIS. THIS PROTEIN COULD FUNCTION AS A TEMPLATE FOR PQQ  
 SYNTHESIS AND PROVIDE THE GLUTAMATE AND TYROSINE RESIDUES WHICH  
 ARE NECESSARY FOR THE SYNTHESIS.  
 CC -1- SIMILARITY: TO OTHER BACTERIAL PQQ.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X06452; NOT ANNOTATED\_CDS.  
 DR PIR: B32252; B32252.  
 KW PQQ.  
 SQ SEQUENCE 24 AA; 2938 MW; 0E671EB9B7EA6838 CRC64;

Query Match 17.9%; Score 26; DB 1; Length 24;  
 Best Local Similarity 40.0%; Pred. No. 8.8e+02;  
 Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 Oy 12 AFSEISVGAE 21  
 Db 7 AFTDLRIQFE 16

RESULT 3  
 ID TPX\_CLOPA STANDARD; PRT; 20 AA.  
 AC P81361;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE PROBABLE THIOLE PEROXIDASE (EC 1.11.1.-) (CP 42) (FRAGMENT).  
 OS Clostridium pasteurianum.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1501;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=W5;  
 RX MEDLINE=98291870; PubMed=9629918;  
 RA Flengsrud R., Skjeldal L.;  
 RT "Two-dimensional gel electrophoresis separation and N-terminal  
 RT sequence analysis of proteins from Clostridium pasteurianum W5.";  
 RL Electrophoresis 19:802-806(1998).  
 CC -1- FUNCTION: HAS ANTIOXIDANT ACTIVITY. COULD REMOVE PEROXIDES OR  
 H(2)O(2) (BY SIMILARITY).  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:  
 4.6, ITS MW IS: 20.2 KDA.  
 CC -1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY. TPX SUBFAMILY.  
 DR INTERPRO: IPR002065; TPX.  
 DR PROSITE: PS01265; TPX; PARTIAL.  
 KW Oxidoreductase; Peroxidase.  
 FT NON\_TER 20  
 SO SEQUENCE 20 AA; 2193 MW; 08178FCD2782E765 CRC64;

Query Match 17.2%; Score 25; DB 1; Length 20;  
 Best Local Similarity 41.7%; Pred. No. 1e+03;

Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 Oy 8 DVSIAFSEISVG 19  
 Db 9 EVTLQKEVKVG 20

RESULT 4  
 ID RML2\_YEAST STANDARD; PRT; 15 AA.  
 AC P36522;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE MITOCHONDRIAL 60S RIBOSOMAL PROTEIN L12 (YML12) (FRAGMENT).  
 GN RMLP12.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=91285106; PubMed=2060626;  
 RA Grohmann L., Graack H.-R., Kruft V., Choll T., Goldschmidt-Reisin S.,  
 RA Kitakawa M.;  
 RT "Extended N-terminal sequencing of proteins of the large ribosomal  
 RT subunit from yeast mitochondria.";  
 RL FEBS Lett. 284:51-56(1991).  
 DR PIR: S17261; S17261.  
 DR SGD: L0002687; RMLP12.  
 KW Ribosomal protein; Mitochondrion.  
 FT NON\_TER 15  
 SO SEQUENCE 15 AA; 1851 MW; 74BCD9FEDDB3900 CRC64;

Query Match 16.6%; Score 24; DB 1; Length 15;  
 Best Local Similarity 36.4%; Pred. No. 1.1e+03;  
 Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 Oy 4 GYHLDVSLAPS 14  
 Db 4 GYFEVIVIRFN 14

RESULT 5  
 ID XYLA\_STRVN STANDARD; PRT; 12 AA.  
 AC P14405;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE XYLOSE ISOMERASE (EC 5.3.1.5) (FRAGMENT).  
 OS Streptomyces violaceoruber.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1935;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=LMG 7183;  
 RX MEDLINE=90104230; PubMed=2604694;  
 RA Vangytsperre W., Ampe C., Kersters-Hilderson H., Tempst P.;  
 RT "Single active-site histidine in D-xylose isomerase from Streptomyces  
 RT violaceoruber. Identification by chemical derivatization and peptide  
 RT mapping.";  
 RL Blochem. J. 263:195-199(1989).  
 CC -1- CATALYTIC ACTIVITY: D-XYLOSE = D-XYLULOSE.  
 CC -1- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.  
 DR HSSP: P37031; IDXL.  
 DR INTERPRO: IPR001998; XYLOSE\_ISOM.  
 DR PROSITE: PS00172; XYLOSE\_ISOMERASE\_1; PARTIAL.

DR PROSITE: PS00173; XYLOSE-ISOMERASE\_2, PARTIAL.  
 KW Isomerase; Pentose shunt; Xylose metabolism; Magnesium.  
 FT NON\_TER 1 1  
 CC ACT\_SITE 5 5  
 FT NON\_TER 12 12  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1375 MW; E749268EB1AAAA1 CRC64;

Query Match 15.9%; Score 23; DB 1; Length 12;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 19 GAFFNKD 26  
 Db 1 GVTFHDD 8

RESULT 6  
 BP37\_LEUMA STANDARD; PRT; 13 AA.  
 AC P81754;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE BRAIN PROTEIN 37F3.  
 OS Leucophaea maderae (Madeira cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
 OC Blaberidae; Blaberidae; Leucophaea.  
 NCBI\_TaxID=6988;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE=Brain;  
 RX MEDLINE=97269266; PubMed-9114447;  
 RA Muren J.E., Naessel D.R.;  
 RT "Seven tachykinin-related peptides isolated from the brain of the  
 RT Madeira cockroach: evidence for tissue-specific expression of  
 RT isoforms.";  
 RL Peptides 18:7-15(1997).  
 CC -1- TISSUE SPECIFICITY: BRAIN.  
 CC -1- MASS SPECTROMETRY: MW=1435.7; METHOD=MALDI.  
 SQ SEQUENCE 13 AA; 1436 MW; 8E532C9DE9A7D1B5 CRC64;

Query Match 15.9%; Score 23; DB 1; Length 13;  
 Best Local Similarity 55.6%; Pred. No. 1.4e+03;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 13 FSEISVGAE 21  
 Db 3 FEESTVSAG 11

RESULT 7  
 AGI\_EUPCH STANDARD; PRT; 18 AA.  
 AC P33888;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE LECTIN (FRAGMENT).  
 OS Euphorbia characias (spurge).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eurosids I; Malpighiales; Euphorbiaceae; Euphorbia.  
 NCBI\_TaxID=3991;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=latex;  
 RX MEDLINE=93357266; PubMed-8353129;  
 RA Stilpe F., Licastro F., Morini M.C., Parente A., Savino G.,  
 RA Abondanza A., Bolognesi A., Falasca A.I., Rossi C.A.;  
 RT "Purification and partial characterization of a mitogenic lectin from

RT the latex of Euphorbia marginata.";  
 RL Biochim. Biophys. Acta 1158:33-39(1993).  
 CC -1- FUNCTION: LECTIN THAT BINDS GALACTOSE, GALACTOSE-CONTAINING SUGARS  
 CC AND GENTIOBIOSIDE. IT IS STRONGLY MITOGENIC FOR HUMAN T LYMPHOCYTES.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- PTM: N-GLYCOSYLATED.  
 CC -1- SIMILARITY: TO E.MARGINATA LECTIN.  
 DR PIR; S36121; S36121.  
 KW Lectin.  
 FT NON\_TER 18 18  
 FT NON\_TER 18 18  
 SQ SEQUENCE 18 AA; 1923 MW; C6F6A1A7B2AB124F CRC64;

Query Match 15.9%; Score 23; DB 1; Length 18;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 GYHLDV 9  
 Db 12 GYXVDV 17

RESULT 8  
 ANFC\_CHICK STANDARD; PRT; 22 AA.  
 ID ANFC\_CHICK  
 AC P21805;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE C-TYPE NATRIURETIC PEPTIDE (CNP).  
 GN NPFC.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Brain;  
 RX MEDLINE=9113166; PubMed=1989595;  
 RA Arimura J.J., Minamino N., Kangawa K., Matsuo H.;  
 RT "Isolation and identification of C-type natriuretic peptide in  
 RT chicken brain.";  
 RL Biochem. Biophys. Res. Commun. 174:142-148(1991).  
 CC -1- FUNCTION: VASORELAXANT ACTIVITY. HAS A CGMP-STIMULATING ACTIVITY.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: BELONGS TO THE NATRIURETIC PEPTIDES FAMILY.  
 DR PIR; JTO581; JTO581.  
 DR InterPro; IPR000663; Natr\_peptide.  
 DR Pfam; PF00212; ANP; 1.  
 DR PRINTS; PR00710; NATPEPTIDES.  
 DR SMART; SM00183; NAT\_PER; 1.  
 DR PROSITE; PS00263; NATRIURETIC\_PEPTIDE; 1.  
 KW Vasoactive.  
 FT DISULFID  
 FT DISULFID 6 22  
 SQ SEQUENCE 22 AA; 2244 MW; FC2A4706D8DAC025 CRC64;

Query Match 15.9%; Score 23; DB 1; Length 22;  
 Best Local Similarity 57.1%; Pred. No. 2.3e+03;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 CLGYHLD 8  
 Db 6 CFGVKLD 12

RESULT 9  
 FMK7\_PSEAE STANDARD; PRT; 23 AA.  
 ID FMK7\_PSEAE  
 AC Q53391;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)

```

DE 15-JUN-1998 (Rel. 36, last annotation update)
DE FTMBRIAL PROTEIN (PILIN) (STRAIN KB7) (FRAGMENT).
GN PIL.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KB7;
RA MEDLINE=94103636; PubMed=7903973;
RX IPATSG;
RT "A multicenter comparison of methods for typing strains of
RT Pseudomonas aeruginosa predominantly from patients with cystic
RT fibrosis. The international Pseudomonas aeruginosa Typing Study
RT Group."
RL J. Infect. Dis. 169:134-142(1994).
RN (2)
RP STRUCTURE BY NMR OF 7-23.
RX STRAIN=KB7;
RC MEDLINE=96110702; PubMed=8845350;
RA Campbell A.P., McInnes C., Hodges R.S., Sykes B.D.;
RT "Comparison of NMR solution structures of the receptor binding
RT domains of Pseudomonas aeruginosa pilI strains PAO, KB7, and PAK:
RT implications for receptor binding and synthetic vaccine design."
RL Biochemistry 34:16255-16268(1995).
CC -1 SUBUNIT: THE PILI ARE POLAR FLEXIBLE FILAMENTS OF ABOUT 5.4
CC NANOMETERS DIAMETER AND 2.5 MICROMETERS AVERAGE LENGTH; THEY
CC CONSIST OF ONLY A SINGLE POLYPEPTIDE CHAIN ARRANGED IN A HELICAL
CC CONFIGURATION OF FIVE SUBUNITS PER TURN IN THE ASSEMBLED PILUS.
CC -1 SIMILARITY: BELONGS TO THE N-ME-PHE PILI FAMILY
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL; S67809; CAB3861.1; -
DR PDB; 1KB7; 29-JAN-96.
DR PDB; 1KB8; 29-JAN-96.
DR InterPro: IPR001082; Pilin.
DR InterPro: IPR001120; Prok_N_methyltn.
DR Pfam: PF00114; pilin; 1.
DR PROSITE: PS00409; PROKAR_NTER_METHYL; PARTIAL.
KW fimbria; 3D-structure.
FT NON_TER
FT DISULFID
SQ SEQUENCE 23 AA; 2415 MW; DOBE77514AF041CF CRC64;

Query Match
Best Local Similarity 15.9%; Score 23; DB 1; Length 23;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 17 SYGAENKDDC 27
DB 11 TVDAKFRPNC 21

RESULT 10
COXC_THUOB
ID COXC_THUOB STANDARD; PRT; 24 AA.
AC P80973;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE VA-2 (EC 1.9.3.1) (FRAGMENT).
OS Thunnus obesus (Bigeye tuna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

```

```

OC Acanthomorpha; Actinopterygii; Percormorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8241;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart, and Liver;
RX MEDLINE=97454291; PubMed=9310366;
RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspelt F.,
RA Kadenbach B.;
RT "The subunit structure of cytochrome-c oxidase from tuna heart and
RT liver."
RL Eur. J. Biochem. 248:99-103(1997).
CC -1 FUNCTION: THIS IS THE HEME A-CONTAINING CHAIN OF CYTOCHROME C
CC OXIDASE, THE TERMINAL OXIDASE IN MITOCHONDRIAL ELECTRON TRANSPORT.
CC -1 CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -1 SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC -1 SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VA FAMILY.
KW Oxidoreductase; Heme; Mitochondrion; Inner membrane.
FT NON_TER
SQ SEQUENCE 24 AA; 2903 MW; 20998FB91F22E43B CRC64;

Query Match
Best Local Similarity 15.9%; Score 23; DB 1; Length 24;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 22 FNKDD 26
DB 19 FNKSD 23

RESULT 11
POQA_PSEFL
ID POQA_PSEFL STANDARD; PRT; 24 AA.
AC P55171;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE COENZYME PQQ SYNTHESIS PROTEIN A.
GN POQA.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHAO;
RX MEDLINE=96064397; PubMed=8526497;
RA Schneider U., Keel C., Defago G., Haas D.;
RT "Tn5-directed cloning of pqq genes from Pseudomonas fluorescens CHAO;
RT mutational inactivation of the genes results in overproduction of the
RT antibiotic pyoverdinin."
RL Appl. Environ. Microbiol. 61:3856-3864(1995).
CC -1 FUNCTION: REQUIRED FOR COENZYME PYROLO-QUINOLINE-QUINONE (PQQ)
CC BIOSYNTHESIS. THIS PROTEIN COULD FUNCTION AS A TEMPLATE FOR PQQ
CC ARE NECESSARY FOR THE SYNTHESIS
CC -1 SIMILARITY: TO OTHER BACTERIAL PQA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL; X87299; CAA60732.1; -
DR EMBL; X87299; CAA60731.1; ALT_INIT.
DR PQQ.
SQ SEQUENCE 24 AA; 2871 MW; ED13A879B2EA3EBE CRC64;

```





```

OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCT 418;
RX MEDLINE=92212283; PubMed=1313537;
RA Meulenber J.J.M., Sellink E., Riegan N.H., Postma P.W.;
RT "Nucleotide sequence and structure of the Klebsiella pneumoniae pqr
  operon.";
RL Mol. Gen. Genet. 232:284-294(1992).
CC -1- FUNCTION: REQUIRED FOR COENZYME PYRROLO-QUINOLINE-QUINONE (PQQ)
CC BIOSYNTHESIS. THIS PROTEIN COULD FUNCTION AS A TEMPLATE FOR PQQ
CC SYNTHESIS AND PROVIDE THE GLUTAMATE AND TYROSINE RESIDUES WHICH
CC ARE NECESSARY FOR THE SYNTHESIS.
CC -1- SIMILARITY: TO OTHER BACTERIAL PQA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X58778; CAA41579.1; .
DR PIR; S20453; S20453.
KW PQQ.
SQ SEQUENCE 23 AA; 2764 MW; ACGB321460871C5D CRC64;

Query Match 15.2%; Score 22; DB 1; Length 23;
Best Local Similarity 40.0%; Pred. No. 3.5e+03;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 12 AFSEISVGA 21
DB 6 AFIDRLGLE 15

RESULT 16
PRTM_SKECO
ID PRTM_SKECO STANDARD; PRT; 25 AA.
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE CYTOCHROME B6-F COMPLEX SUBUNIT VII (CYTOCHROME B6F COMPLEX SUBUNIT
DE PPTM) (FRAGMENT).
GN PPTM OR YCF31.
OS Skeletomena costatum.
OC Chloroplast.
OC Eukaryota; Stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Thalassiosiristhycidae; Thalassiosirales; Skeletonemataceae;
OX NCBI_TaxID=2843;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIES-323;
RA Tada N., Otsuka S., Oyaizu H., Matsumoto S.;
RT "Plastid DNA sequences of Skeletomena costatum NIES 323.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CYTOCHROME B6-F COMPLEX SUBUNIT.
CC -1- SUBCELLULAR LOCATION: THYLAKOID MEMBRANE-ASSOCIATED (BY
CC SIMILARITY)
CC -1- SIMILARITY: BELONGS TO THE PPTM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL; AJ132266; CAA10631.1; .
KW Electron transport; Chloroplast; Respiratory chain; Thylakoid;
KW Transmembrane.
FT NON_TER 1
FT TRANSMEM 2 22
SQ SEQUENCE 25 AA; 2630 MW; 99ECF1B6CA3113AA CRC64;

Query Match 15.2%; Score 22; DB 1; Length 25;
Best Local Similarity 38.5%; Pred. No. 3.8e+03;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 9 VSLAFSEISVGA 21
DB 13 LSLGFAFLVQGE 25

RESULT 17
CX7A_CONTE
ID CX7A_CONTE STANDARD; PRT; 27 AA.
AC P24160;
DT 01-MAR-1992 (Rel. 21, Created)
DT 30-MAY-2000 (Rel. 38, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CONOTOXIN TXVIIA (TXIIA).
OS Conus textile (Cloth-of-gold cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conidae; Conus.
OX NCBI_TaxID=6494;
RN [1]
RP SEQUENCE.
RC STRAIN=NEOVICARIUS; TISSUE=Venom;
RX MEDLINE=92104183; PubMed=1761058;
RA Fainzilber M., Gordon D., Hesson A., Spira M.E., Zlotkin E.;
RT "Mollusc-specific toxins from the venom of Conus textile
RT neovicarius.";
RL Eur. J. Biochem. 202:589-595(1991).
RN [2]
RP REVISIONS TO 1 AND C-TERMINAL, AND MASS-SPECTROMETRY.
RX MEDLINE=97022130; PubMed=8868490;
RA Nakamura T., Yu Z., Fainzilber M., Burlingame A.L.;
RT "Mass spectrometric-based revision of the structure of a cysteine-rich
RT peptide toxin with gamma-carboxyglutamic acid, TXVIIA, from the sea
RT snail, Conus textile.";
RL Protein Sci. 5:524-530(1996).
CC -1- FUNCTION: POTENT NEUROTOXIN. MAY EXERT ITS EFFECTS AT THE LEVEL
CC OF THE NEUROMUSCULAR JUNCTION.
CC -1- PTM: THREE DISULFIDE BONDS ARE PRESENT.
CC -1- MASS SPECTROMETRY: MW=3088.9; METHOD=ELECTROSPRAY.
DR PIR; S19620; S19620.
KW Venom; Neurotoxin; Vitamin K; Gamma-carboxyglutamic acid; Amidation.
FT MOD_RES 9 9
FT MOD_RES 13 13 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 27 27 AMIDATION.
SQ SEQUENCE 27 AA; 3008 MW; D7A49781300FE1E7 CRC64;

Query Match 14.8%; Score 21.5; DB 1; Length 27;
Best Local Similarity 26.9%; Pred. No. 4.8e+03;
Matches 7; Conservative 3; Mismatches 9; Indels 7; Gaps 1;

QY 2 CIGYHLDVSLAFSEISVGAEFNKDC 27
DB 1 CGGYS-----TYCEVDSECCSDNC 19

RESULT 18
TKL4_LOCM1
ID TKL4_LOCM1 STANDARD; PRT; 10 AA.
AC P30250;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)

```

DT 01-NOV-1995 (Rel. 32, last annotation update)  
 DE LOCUSTATACHYKININ IV (TK-IV).  
 OS Locusta migratoria (Migratory locust).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthoptera; Orthopteroidea; Orthoptera; Caelifera;  
 OC Acridomorpha; Acridoidea; Acrididae; Locusta.  
 OX NCBI\_TaxID=7004;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Brain;  
 RX MEDLINE=91219696; PubMed=2132575;  
 RA Schoofs L., Holman G.M., Hayes T.K., Kochansky J.P., Nachman R.J.,  
 de Loof A.;  
 RT "Locustatachykinin III and IV: two additional insect neuropeptides  
 with homology to peptides of the vertebrate tachykinin family.";  
 RL Regul. Pept. 31:199-212(1990).  
 CC -1- FUNCTION: MYOACTIVE PEPTIDE. STIMULATES THE CONTRACTION OF THE  
 OVIDUCT AND FOREGUT.  
 CC -1- SIMILARITY: SOME SIMILARITY TO TACHYKININS.  
 CC PIR: B60073; ECIQ4M.  
 DR Tachykinin; Neuropeptide; Amladation.  
 KW MOD.RES 10  
 FT SEQUENCE 10 AA; 1040 MW; 9E52CD71E9C87735 CRC64;

Query Match 14.5%; Score 21; DB 1; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 2.1e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGYH 6  
 1111  
 DB 4 LGFH 7

RESULT 19  
 FIBA\_SHEEP STANDARD; PRT; 19 AA.  
 AC P14451;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, last sequence update)  
 DT 01-MAY-1991 (Rel. 18, last annotation update)  
 DE FIBRINOPEPTIDE A.  
 OS Ovis aries (Sheep), and Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940, 9925;  
 RN [1]  
 RP SEQUENCE.  
 RA Blomback B., Grondahl N.J.;  
 RT "Studies on fibrinopeptides from mammals.";  
 RL Acta Chem. Scand. 19:1789-1791(1965).  
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 AGGREGATION.  
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 CC HSSP: P02671; 1BBR.  
 DR Blood coagulation; Plasma.  
 KM NON\_TER 19  
 FT SEQUENCE 19 AA; 1848 MW; 99C98873D1B059C5 CRC64;

Query Match 14.5%; Score 21; DB 1; Length 19;  
 Best Local Similarity 80.0%; Pred. No. 4.1e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 18 VGEF 22  
 1111

DB 7 VGEF 11

RESULT 20  
 V101\_VACCV STANDARD; PRT; 20 AA.  
 AC P16714;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, last sequence update)  
 DT 01-FEB-1991 (Rel. 17, last annotation update)  
 DE PROTEIN I1 (FRAGMENT).  
 GN I1.  
 OS Vaccinia virus (strain WR).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Orthopoxvirus.  
 OX NCBI\_TaxID=10254;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88215015; PubMed=2835495;  
 RA Schmitt J.F.C., Stuenkelberg H.G.;  
 RT "Sequence and transcriptional analysis of the vaccinia virus HindIII  
 I fragment";  
 RL J. Virol. 62:1889-1897(1988).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: J03399; AAB59803.1; -.  
 DR PIR: A29889; A29889.  
 KW Late protein.  
 KM NON\_TER 20  
 FT SEQUENCE 20 AA; 2241 MW; 6E97AA78D505C4B1 CRC64;

Query Match 14.5%; Score 21; DB 1; Length 20;  
 Best Local Similarity 50.0%; Pred. No. 4.3e+03;  
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 11 LAFSEISVGA 20  
 11111111  
 DB 8 LVFNSISARA 17

RESULT 21  
 RL10\_KLEPN STANDARD; PRT; 22 AA.  
 AC P41190;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, last sequence update)  
 DT 20-AUG-2001 (Rel. 40, last annotation update)  
 DE 50S RIBOSOMAL PROTEIN L10 (FRAGMENT).  
 GN RPLJ.  
 OS Klebsiella pneumoniae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Klebsiella.  
 OX NCBI\_TaxID=573;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCTC 5054;  
 RA Znyvoloup A.N.;  
 RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X74445; CAA52456.1; -

DR PIR: S35975; S35975.

DR PIR: S35976; S35976.

DR InterPro: IPR002363; Ribosomal\_L10\_eub.

DR PROSITE: PS01109; RIBOSOMAL\_L10; PARTIAL.

KW Ribosomal protein.

FT INIT\_MET 0

FT NON\_TER 22

SEQUENCE 22 AA; 2268 MW; 113D96520A02CD78 CRC64;

Query Match

Best Local Similarity 14.5%; Score 21; DB 1; Length 22;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 14 SEISVGA 20

DB 15 SEVAKGA 21

RESULT 22

RL10\_ENTCL

ID RL10\_ENTCL STANDARD; PRT; 23 AA.

AC 047608; 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE 50S RIBOSOMAL PROTEIN L10 (FRAGMENT).

GN RPLJ.

OS Enterobacter cloacae.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Enterobacter.

OX NCBI\_TaxID=550;

RN [1]

RP SEQUENCE FROM N.A.

RA Zhyvoloup A.;

RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: X74444; CAA52455.1; -

DR InterPro: IPR002363; Ribosomal\_L10\_eub.

DR PROSITE: PS01109; RIBOSOMAL\_L10; PARTIAL.

KW Ribosomal protein.

FT INIT\_MET 0

FT NON\_TER 23

SEQUENCE 23 AA; 2355 MW; 39413D96520A02CD CRC64;

Query Match

Best Local Similarity 14.5%; Score 21; DB 1; Length 23;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 14 SEISVGA 20

DB 15 SEVAKGA 21

RESULT 23

CAMT\_PINPS STANDARD; PRT; 24 AA.

ID CAMT\_PINPS STANDARD; PRT; 24 AA.

AC P81081; 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE PROBABLE CAFFEoyl-COA O-METHYLTRANSFERASE (EC 2.1.1.104) (TRANS-

DE CAFFEoyl-COA 3-O-METHYLTRANSFERASE) (CCOAMT) (CCOAMT) (WATER STRESS

DE RESPONSIVE PROTEIN 13) (FRAGMENTS).

OS Pinus pinaster (Maritime pine).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Coniferales; Pinaceae; Pinus.

OX NCBI\_TaxID=71647;

RN [1]

RP SEQUENCE.

RC TISSUE=Needle;

RX MEDLINE=98418576; PubMed=9747804;

RA Costa P., Bahman N., Frigerio J.-M., Kremer A., Plomion C.;

RT "Water-deficit-responsive proteins in maritime pine.";

RL Plant Mol. Biol. 38:587-596(1998).

RP SEQUENCE.

RC TISSUE=Needle;

RX MEDLINE=99274088; PubMed=10344291;

RA Costa P., Plomion C., Bauw G., Dubos C., Bahman N., Kremer A.,

RT Frigerio J.-M., Plomion C.;

RT "Separation and characterization of needle and xylem maritime pine

RT proteins.";

RL Electrophoresis 20:1098-1108(1999).

CC -1- FUNCTION: INVOLVED IN THE REINFORCEMENT OF THE PLANT CELL WALL

CC SIMILARITY).

CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + CAFFEoyl-COA =

CC -1- S-ADENOSYL-L-HOMOCYSTEINE + FERULOYL-COA.

CC -1- PATHWAY: LIGNIN BIOSYNTHESIS.

CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN

CC IS: 5.2. ITS MW IS: 30 KDA.

KW Lignin biosynthesis; Transferase; Methyltransferase.

FT NON\_TER 1

FT NON\_CONS 9

FT NON\_TER 24

SEQUENCE 24 AA; 2447 MW; AABE842F9EB0CF0 CRC64;

Query Match

Best Local Similarity 14.5%; Score 21; DB 1; Length 24;

Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 10 SLAFSEISVG 19

DB 9 NIEISQIPVG 18

RESULT 24

ANF\_ANGJA

ID ANF\_ANGJA STANDARD; PRT; 27 AA.

AC P18144; 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE ATRIAL NATRIURETIC FACTOR (ANF) (ATRIAL NATRIURETIC PEPTIDE) (ANP).

OS Anguilla japonica (Japanese eel).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;

OC Anguillidae; Anguilla.

OX NCBI\_TaxID=7937;

RN [1]

RP SEQUENCE.

RC TISSUE=Heart atrium;

RX MEDLINE=90026430; PubMed=2529857;

RA Takei Y., Takahashi A., Watanabe T.X., Nakajima K., Sakakibara S.;

RT "Amino acid sequence and relative biological activity of eel atrial

RT natriuretic peptide.";

RL Biochem. Biophys. Res. Commun. 164:537-543(1989).

CC -1- FUNCTION: EXHIBITS NATRIURETIC AND VASODEPRESSOR ACTIVITY. HAS A

CC CGMP-STIMULATING ACTIVITY.

CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: BELONGS TO THE NATRIURETIC PEPTIDES FAMILY.  
 DR PIR; A33431; A33431.  
 DR InterPro: IPR000663; Natr\_peptide.  
 DR Pfam: PF00212; ANP; 1.  
 DR PRINTS; PR00710; NATPEPTIDES.  
 DR SMART; SM00183; NAT\_PEP; 1.  
 DR PROSITE; PS00263; NATRIURETIC\_PEPTIDE; 1.  
 KW Vasoactive.  
 FT DISULFID 7 23  
 SQ SEQUENCE 27 AA; 2792 MW; 9EF510CCF282B4D CRC64;

Query Match 14.5%; Score 21; DB 1; Length 27;  
 Best Local Similarity 57.1%; Pred. No. 5.8e+03;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CLGYHLD 8  
 DB 7 CFSGKLD 13

RESULT 25  
 CD59\_SHEEP  
 ID CD59\_SHEEP STANDARD; PRT; 27 AA.  
 AC P58020;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE CD59 GLYCOPROTEIN (MEMBRANE ATTACK COMPLEX INHIBITION FACTOR) (MACIF)  
 DE (MAC-INHIBITORY PROTEIN) (MAC-IP) (PROTECTIN) (COMPLEMENT INHIBITORY  
 DE PROTEIN) (CIP) (FRAGMENT).  
 GN CD59.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Erythrocyte;  
 RX MEDLINE=93239201; PubMed=7682985;  
 RA van den Berg C.W., Harrison R.A., Morgan B.P.;  
 RT "The sheep analogue of human CD59: purification and characterization  
 of its complement inhibitory activity.";  
 RL Immunology 78:349-357(1993).  
 CC -1- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK  
 CC COMPLEX (MAC) ACTION. ACTS BY BINDING, IN A SPECIES SPECIFIC  
 CC MANNER, TO THE C8 AND/OR C9 COMPLEMENTS OF THE ASSEMBLING MAC,  
 CC THEREBY PREVENTING INCORPORATION OF THE MULTIPLE COPIES OF C9  
 CC REQUIRED FOR COMPLETE FORMATION OF THE OSMOTIC PORE.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ERYTHROCYTES AND LYMPHOCYTES.  
 CC NOT DETECTED IN PLATELETS.  
 DR PROSITE; PS00983; LY6\_UPAR; PARTIAL.  
 KW Antigen; Glycoprotein.  
 FT NON\_TER 1 1  
 FT DISULFID 3 25  
 FT DISULFID 6 12  
 FT CARBOHYD 17 17  
 FT UNSURE 3 3  
 FT UNSURE 6 6  
 FT UNSURE 17 17  
 FT UNSURE 17 17  
 FT UNSURE 18 18  
 FT UNSURE 25 25  
 FT NON\_TER 27 27  
 SQ SEQUENCE 27 AA; 3025 MW; 92E6C978EC7663FE CRC64;

Query Match 14.5%; Score 21; DB 1; Length 27;  
 Best Local Similarity 30.0%; Pred. No. 5.8e+03;  
 Matches 3; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 18 VGAEFNKDDC 27  
 DB 16 INCTXNODAC 25

Search completed: January 6, 2002, 09:17:45  
 Job time: 278 sec

**THIS PAGE BLANK (USPTO)**







```
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TEMBLrel. 14, Last annotation update)
DE HAPTOGLOBIN-RELATED PROTEIN BETA SUBUNIT (FRAGMENT).
OC Trypanosoma brucei.
OC Eukaryote; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
RN NCBI_TaxID=5691;
RX MEDLINE=95232503; Pubmed=7716520;
RA Smith A.B., Esko J.D., Hajduk S.L.;
RT "Killing of trypanosomes by the human haptoglobin-related protein.";
RL Science 268:284-286(1995).
SQ SEQUENCE 19 AA; 2055 MW; 65BD135667C94056 CRC64;

Query Match
Best Local Similarity 19.3%; Score 28; DB 5; Length 19;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 LGYHLDVSLAF 13
Db 2 LGHLDKSGSF 12

RESULT 7
Q82392 PRELIMINARY; PRT; 21 AA.
AC 082392;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
DE T-CELL LYMPHOTROPIC VIRUS TYPE 2 TAX (FRAGMENT).
GN TAX.
OS Human T-cell leukemia virus type II (HTLV-II).
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11909;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96190544; Pubmed=8627666;
RA Eitraku N., Novoa P., Monken C., Oliviera M.D., Oliviera O.D.,
RA Ishak R., Oliviera M.P., Laureiro P., Ishak M., Acevedo V.,
RA Hammershlag N., Zhu S.W., Kubo T., Hall W.W.;
RT "Identification and characterization of a new and distinct molecular
RT subtype of human T-cell lymphotropic virus type 2.";
RL J. Virol. 70:1481-1492(1996).
DR EMBL: U32871; AAB04905.1; -.
FT NON_TER 1
SQ SEQUENCE 21 AA; 2383 MW; 11F39408575CF6D0 CRC64;

Query Match
Best Local Similarity 19.3%; Score 28; DB 12; Length 21;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 13 FSEISVGAERFKDD 26
Db 2 YTNIPVSILFNKEE 15

RESULT 8
Q82410 PRELIMINARY; PRT; 21 AA.
AC 082410;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
DE T-CELL LYMPHOTROPIC VIRUS TYPE 2 TAX (FRAGMENT).
GN TAX.
OS Human T-cell leukemia virus type II (HTLV-II).
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11909;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=96190544; Pubmed=8627666;
RA Eitraku N., Novoa P., Monken C., Oliviera M.D., Oliviera O.D.,
RA Ishak R., Oliviera M.P., Laureiro P., Ishak M., Acevedo V.,
RA Hammershlag N., Zhu S.W., Kubo T., Hall W.W.;
RT "Identification and characterization of a new and distinct molecular
RT subtype of human T-cell lymphotropic virus type 2.";
RL J. Virol. 70:1481-1492(1996).
DR EMBL: U32883; AAB04923.1; -.
FT NON_TER 1
SQ SEQUENCE 21 AA; 2383 MW; 11F39408575CF6D0 CRC64;

Query Match
Best Local Similarity 19.3%; Score 28; DB 12; Length 21;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 13 FSEISVGAERFKDD 26
Db 2 YTNIPVSILFNKEE 15

RESULT 9
Q82411 PRELIMINARY; PRT; 21 AA.
AC 082411;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
DE T-CELL LYMPHOTROPIC VIRUS TYPE 2 TAX (FRAGMENT).
GN TAX.
OS Human T-cell leukemia virus type II (HTLV-II).
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11909;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96190544; Pubmed=8627666;
RA Eitraku N., Novoa P., Monken C., Oliviera M.D., Oliviera O.D.,
RA Ishak R., Oliviera M.P., Laureiro P., Ishak M., Acevedo V.,
RA Hammershlag N., Zhu S.W., Kubo T., Hall W.W.;
RT "Identification and characterization of a new and distinct molecular
RT subtype of human T-cell lymphotropic virus type 2.";
RL J. Virol. 70:1481-1492(1996).
DR EMBL: U32884; AAB04924.1; -.
FT NON_TER 1
SQ SEQUENCE 21 AA; 2383 MW; 11F39408575CF6D0 CRC64;

Query Match
Best Local Similarity 19.3%; Score 28; DB 12; Length 21;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 13 FSEISVGAERFKDD 26
Db 2 YTNIPVSILFNKEE 15

RESULT 10
Q82412 PRELIMINARY; PRT; 21 AA.
AC 082412;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
DE T-CELL LYMPHOTROPIC VIRUS TYPE 2 TAX (FRAGMENT).
GN TAX.
OS Human T-cell leukemia virus type II (HTLV-II).
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11909;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96190544; Pubmed=8627666;
RA Eitraku N., Novoa P., Monken C., Oliviera M.D., Oliviera O.D.,
```



Db 1 MSFRINTNIALTSHAVGVONNRD 24

## RESULT 15

09DE23 PRELIMINARY; PRT; 20 AA.  
AC 09DE23;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE UORF2.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=20528616; PubMed=11073974;  
RX Kobayashi M., Yu R.T., Yasuda K., Umesono K.;  
RT "Cell-type-specific regulation of the retinoic acid receptor mediated  
RT by the orphan nuclear receptor TLX";  
RL Mol. Cell. Biol. 20:8731-8739(2000).  
DR EMBL; AF220160; AAG35363.1; -.  
DR InterPro; IPR000345; Cytochrome\_heme\_bind.  
DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
SO SEQUENCE 20 AA; 2261 MW; AA7738B0BE3C482 CRC64;

Query Match 17.9%; Score 26; DB 13; Length 20;  
Best Local Similarity 66.7%; Pred. No. 3e+03;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 6 HLDVSLAFSEISVG 14  
Db 5 HLDQLQAFS 13

## RESULT 16

063076 PRELIMINARY; PRT; 21 AA.  
AC 063076;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE RAT ALPHA-2 TYPE I COLLAGEN (FRAGMENT).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=85122694; PubMed=6395893;  
RX Genovese C., Rowe D., Kream B.;  
RT "Construction of DNA sequences complementary to rat alpha 1 and alpha  
RT 2 collagen mRNA and their use in studying the regulation of type I  
RT collagen synthesis by 1,25-dihydroxyvitamin D.";  
RL Biochemistry 23:6210-6216(1984).  
DR EMBL; M12200; AAA40835.1; -.  
DR InterPro; IPR000885; FIB\_collagen\_C.  
DR Pfam; PF01410; COLFI; 1.  
FT NON\_TER 1 21  
FT NON\_TER 1 21  
SO SEQUENCE 21 AA; 2358 MW; B86A6ECCEEA6D60 CRC64;

Query Match 17.9%; Score 26; DB 11; Length 21;  
Best Local Similarity 33.3%; Pred. No. 3.1e+03;  
Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 5 YHLDVSLAFSEISVG 19  
Db 7 YHCKNSIAYLDEETG 21

## RESULT 17

09Z274 PRELIMINARY; PRT; 22 AA.  
AC 09Z274;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE NADH1 (FRAGMENT).  
OS Phlebotomus bergeroti.  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Psychodidae;  
OC Psychodidae; Phlebotomus.  
OX NCBI\_TaxID=59273;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Essegbir E., Ready P.D., Killick-Kendrick R., Ben Ismail R.;  
RT "Mitochondrial haplotypes and phylogeography of Phlebotomus vectors of  
RT Leishmania major";  
RL Insect Mol. Biol. 6:0-0(1997).  
DR EMBL; U94474; AAD09262.1; -.  
DR Mitochondrion.  
KW Mitochondrion.  
FT NON\_TER 1 1  
SO SEQUENCE 22 AA; 2453 MW; CD5B7EDC1A37035A CRC64;

Query Match 17.9%; Score 26; DB 8; Length 22;  
Best Local Similarity 35.7%; Pred. No. 3.3e+03;  
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 6 HLDVSLAFSEISVG 19  
Db 3 YLPASLCFISLFTG 16

## RESULT 18

09Z272 PRELIMINARY; PRT; 22 AA.  
AC 09Z272;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE NADH1 (FRAGMENT).  
OS Phlebotomus papatasi.  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Psychodidae;  
OC Psychodidae; Phlebotomus.  
OX NCBI\_TaxID=29031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Essegbir E., Ready P.D., Killick-Kendrick R., Ben Ismail R.;  
RT "Mitochondrial haplotypes and phylogeography of Phlebotomus vectors of  
RT Leishmania major";  
RL Insect Mol. Biol. 6:0-0(1997).  
DR EMBL; U94475; AAD09264.1; -.  
DR Mitochondrion.  
KW Mitochondrion.  
FT NON\_TER 1 1  
FT NON\_TER 1 1  
SO SEQUENCE 22 AA; 2453 MW; CD5B7EDC1A37035A CRC64;

Query Match 17.9%; Score 26; DB 8; Length 22;  
Best Local Similarity 35.7%; Pred. No. 3.3e+03;  
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 6 HLDVSLAFSEISVG 19  
Db 3 YLPASLCFISLFTG 16

## RESULT 19

09Q0123		
ID	09Q0123	PRELIMINARY;            PRT;            27 AA.
AC	09Q0123;	
DT	01-MAY-2000	(TREMBLrel, 13, Created)
DT	01-MAY-2000	(TREMBLrel, 13, Last sequence update)
DE	01-JUN-2001	(TREMBLrel, 17, Last annotation update)
DE	GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2/NSI (GP68)]	
DE	[FRAGMENT].	
OS	Hepatitis C virus.	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;	
OX	Hepadnaviruses.	
NCBI_Taxid=11103;		
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Saadtes K., Dubois M., Pasquier C., Izopet J.;	
RT	"The genetic heterogeneity of hepataviral region 1 of the viral	
RT	genome and the sensitivity of hepatitis C virus to interferon alpha	
RT	therapy.";	
RL	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.	
-1-	SUBUNIT: HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E2/NSI IS	
CC	CONTAINED WITHIN THE GENOME POLYPEPTIDE THAT CONTAINS: CAPSID	
CC	PROTEIN C; THE ENVELOPE GLYCOPROTEIN E1; ENVELOPE GLYCOPROTEIN	
CC	E2/NSI; NONSTRUCTURAL PROTEINS NS2, NS3, NS4A, NS4B, NS5A AND	
CC	NS5B. THE VIRION OF THE VIRUS IS COVERED BY A LIPIDPROTEIN ENVEL	
CC	THAT CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E (BY	
CC	STABILITY).	
-1-	SMILARITY: TO HEPATITIS C VIRUS NON-STRUCTURAL PROTEIN E2/NSI	
CC	FAMILY.	
DR	EMBL: AF166754; AAD52412.1; -.	
DR	InterPro: IPR002531; HCV_NSI.	
DR	Pfam: PF01560; HCV_NSI; 1.	
KW	Coat protein; Envelope protein; Glycoprotein; Nonstructural protein	
KW	Polyprotein; Transmembrane.	
FT	NON_TER	1
FT	NON_TER	27
SO	SEQUENCE	27 AA; 2778 MM; 9D46112E115CA85A CRC64;

Query Match	Best Local Similarity	17.98;	Score 26;	DB 12;	Length 27;
Matches 8;	Conservative	53.3%;	Pred. No. 4.1e+03;	Mismatches 6;	Indels 0;
0;	Gaps	0;			
QY	9 VSLAFSEISVGAEFN 23	11 1 1111: 1			
DB	13 VSKFTSLFVGAQON 27				
RESULT	20				
092820					
ID	092820	PRELIMINARY;	PRT;	13 AA.	
DT	092820;				
DC	01-NOV-1998 (TREMBLrel. 08, Created)				
DT	01-NOV-1998 (TREMBLrel. 08, last sequence update)				
DE	01-NOV-1998 (TREMBLrel. 08, last annotation update)				
DE	ENVELOPE GLYCOPROTEIN (FRAGMENT).				
GN	ENV.				
OS	Chimpanzee immunodeficiency virus (SIV/cpz2) (CIV).				
OC	Viruses; Retroid Viruses; Retroviridae; Lentivirus.				
OX	NCBI_Taxid=11723;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=SIIVSMB7;				
RA	MEDLINE=96171634; Pubmed=8574147;				
RA	Kraiselburg E.N., Torres J.V.;				
RT	"Properties of virus-like particles produced by SIV-chronically				
RT	infected human cell clones."				
Cell.	MOL. Biol. 41:41-52(1995).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=SIIVSMB7;				
RA	MEDLINE=98109578; Pubmed=9449524;				
RA	Kraiselburg E.N., Salaman A., Baltran M., Rivera M., Oliver J.,				
RA	Kessler M., Rostovich A., Rodriguez A., Biliska M., Montefiori D.,				

RA	Torres-Bauza L.J., Martinez I., "Vaccine evaluation studies of replication-defective SIVsmB7.", Cell. Mol. Biol. 43:915-924(1997).			
RL	EMBL; AF034690; AAC09235.1; -.			
KR	Envelope protein..			
FT	NON_TER 13 13			
SQ	SEQUENCE 13 AA; 1359 MW; 06E20308931E36D4 CRC64;			
Query Match				
Best Local Similarity	17.28;	Score 25;	DB 12;	Length 13;
Matches 5; Conservative	50.0%;	Pred. No. 2.6e+03;	Mismatches 3;	Indels 0;
Gaps				0.
OY	2 CIGHYLDVSL 11      :::			
Dd	3 CLGNQLIAL 12			

RESULT	21		
09RS57			
ID	Q9RS57	PRELIMINARY;	PRT; 16 AA.
AC	Q9RS57;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)		
DE	OUTER SHEATH PROTEIN.		
OS	Treponema denticola.		
OC	Bacteria; Spirochaetales; Spirochaetaceae; Treponema.		
OX	NCBI_TaxID=158;		
RN	[1]		
RP	SEQUENCE.		
RX	MEDLINE=92041580; PubMed=1938897;		
RA	Weinberg A., Holt S.C.;		
RT	"Chemical and biological activities of a 64-kilodalton outer sheath		
RT	protein from <i>Treponema denticola</i> strains."		
RL	J. Bacteriol. 173:6935-6947(1991).		
SO	SEQUENCE 16 AA; 1640 MW; 3245DE6C11C5D124 CRC64;		

Query Matched	17.2%	Score 25;	DB 2;	Length 16;
Best Local Similarity	35.7%	Pred. No. 3.3e+03;		
Matches 5;	Conservative	3;	Mismatches 6;	Indels 0;
			Gaps	0;
OY	7 LDVSLAFSEISVGA 20			
Db	2 LDIALDLGEAXINA 15			
RESULT 22				
061946				
ID 061946	PRELIMINARY;	PRT;	24 AA.	
AC 061946;				
DT 01-NOV-1996 (TREMBLrel. 01, Created)				
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)				
DT 01-NOV-1999 (TREMBLrel. 12, last annotation update)				
DE NEURAL CELL ADHESION MOLECULE (FRAGMENT).				
GN NCAM.				
OS Mus musculus (Mouse).				
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX NCBI_TaxID=10090;				
RN (1)				
RP SEQUENCE FROM N.A.				
RP MEDLINE=88283628; PubMed=3396534;				
RA Barbas J.A., Chaix J.C., Steinmetz M., Goridis C.;				
RT "Differential splicing and alternative polyadenylation generates				
RT distinct NCAM transcripts and proteins in the mouse.";				
RL EMBL, X07196; CAA30174.1; -.				
DR MGD; MGI:97281; Ncam				
DR NON_TER				
FT				
SO SEQUENCE	24 AA;	2457 MW;	D529FAB2495D02D9	CRC64;



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:05:07 ; Search time 22.59 Seconds  
(without alignments)  
88.534 Million cell updates/sec

Title: US-09-020-393b-18

Perfect score: 145  
Sequence: 1 KCLGYHLDVSLAFSEISVGAFFNKDDC 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 227461

Minimum DB seq length: 0  
Maximum DB seq length: 27

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq\_1101.\*  
1: /SID52/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	145	100.0	27	20	AAV27328
2	35	24.1	15	21	AAV77161
3	34	23.4	19	22	AAE00613
4	34	23.4	20	18	AAW35697
5	34	23.4	21	19	AAW82186
6	34	23.4	21	22	AAW33203
7	34	23.4	21	22	AAW633204
8	34	23.4	26	21	AAW96324
9	34	23.4	26	22	AAW70374
10	34	23.4	27	21	AAW37016
11	33	22.8	13	22	AAU15328

12	33	22.8	13	22	AAU15338	Schizophrenia-asso
13	33	22.8	15	19	AAW38432	Human dendritic ce
14	33	22.8	16	21	AAW06316	Human beta-amyloid
15	33	22.8	18	22	AAE00609	Beta-amyloid precu
16	33	22.8	19	22	AAE00611	Amyloid precursor
17	33	22.8	20	21	AAW69714	Beta-APP alpha-sec
18	33	22.8	20	21	AAW69716	Beta-APP alpha-sec
19	33	22.8	21	18	AAW08361	Beta-secretase sub
20	33	22.8	21	19	AAW82188	Fluorogenic protea
21	33	22.8	21	20	AAW33755	Synthetic oligopep
22	33	22.8	21	22	AAW73205	Protease binding s
23	33	22.8	21	22	AAW73206	Protease binding s
24	33	22.8	21	22	AAW73316	Protease indicator
25	33	22.8	21	22	AAW47265	Peptide 5-16'/SW, f
26	32	22.1	10	20	AAW47020	Immunogenic peptid
27	32	22.1	10	22	AAW43201	Mycoplasma genital
28	32	22.1	10	22	AAW43203	Mycoplasma genital
29	32	22.1	10	22	AAW43205	Mycoplasma genital
30	32	22.1	11	21	AAW54223	HLA binding peptid
31	32	22.1	16	21	AAW06315	Human beta-amyloid
32	32	22.1	16	21	AAW06317	Human beta-amyloid
33	32	22.1	18	22	AAE00608	Beta-amyloid precu
34	32	22.1	20	16	AAW64641	HPF3 peptide deriv
35	32	22.1	20	21	AAW69713	Beta-APP alpha-sec
36	32	22.1	20	21	AAW69715	Beta-APP alpha-sec
37	32	22.1	20	22	AAW55243	Anti-HPV3 F1 regio
38	32	22.1	21	16	AAW64642	HPF3 peptide deriv
39	32	22.1	21	22	AAW55244	Anti-HPV3 F1 regio
40	32	22.1	22	16	AAW64643	HPF3 peptide deriv
41	32	22.1	22	22	AAW55245	Anti-HPV3 F1 regio
42	32	22.1	23	16	AAW64644	HPF3 peptide deriv
43	32	22.1	23	22	AAW55246	Anti-HPV3 F1 regio
44	32	22.1	24	16	AAW64645	HPF3 peptide deriv
45	32	22.1	24	22	AAW55247	Anti-HPV3 F1 regio

#### ALIGNMENTS

RESULT 1	AAV27328	standard; peptide: 27 AA.
XX	AAV27328;	
AC		
XX		
DT	05-NOV-1999	(first entry)
XX		
DE	Human C9 protein fragment (residues 359-384).	
XX		
KW	CD59 mediated complement; human; C9 protein; mimetic;	
KW	tumour therapy; complement-mediated inflammation; immune disorder;	
KW	Immunovascularitis; rheumatoid arthritis; scleroderma; C5b-9 complex.	
XX		
OS	Homo sapiens.	
XX		
PN	WO940115-A2.	
XX		
PD	12-AUG-1999.	
XX		
PF	09-FEB-1999;	99WO-US02802.
XX		
PR	09-FEB-1998;	98US-0020393.
XX		
PA	(BLOO-) BLOOD CENT RES FOUND INC.	
PA	(OKLA-) OKLAHOMA MEDICAL RES FOUND.	
XX		
PI	Sims PJ;	
XX		
DR	WPI; 1999-527301/44.	
XX		
PT	Compounds modulating CD59 mediated complement activity, for	
PT	treatment of, e.g. immunovascularitis	
XX		

PS Example 2; Page 36; 75pp; English.  
XX

The invention relates to compounds modulating CD59 mediated complement activity. It provides (i) molecules structurally mimicking human CD59 amino acid residues 42-58 /region which serves as binding site for CD59 C9 interactions) when they are in a spatial orientation which can inhibit the formation of the human C5b-9 complex. These mimetics specifically bind to amino acid residues 359-384 of human C9. (ii) molecules structurally mimicking C9 amino acids 359-384 when they are in a spatial orientation which can promote the formation of the C5b-9 complex. Compounds that mimic CD59 can be used to increase CD59 inhibition of C5b-9 complex assembly. This is especially useful in patients in need of suppression of complement-mediated inflammation, e.g., in immune disorders and diseases such as immunosculitis, rheumatoid arthritis, scleroderma. Compounds that mimic C9 can be used to promote C5b-9 complex assembly. This is useful in patients in need of complement activation. The composition can be administered as an adjunct to tumour therapy. The present sequence represents a human C9 protein fragment.

Query Match	100.0%;	Score 145;	DB 20;	Length 27;
Best Local Similarity	100.0%;	Pred. No. 3,4e-17;		
Matches 27;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

1 KCLGYHLDVSLAESEISVGAEFNKDDC 27  
1 KCLGYHLDVSLAESEISVGAEFNKDDC 27

```

RESULT      2
ID          AAY77161
            standard; peptide; 15 AA.

```

08-MAY-2000 (first entry)

FGFR-binding peptide 13-1, SEQ ID NO:23.

W Fibroblast growth factor receptor; EGF ligand; activation,  
W phage display library; agonist; antagonist; drug screening;  
W competitive inhibitor; angiogenic; wound healing.

Synthetic.

WO200003245-A1

20-JAN-2000.

28-MAY-1999; 99WO-US11844.

28-MAY-1998; 98US-0087107.

(CHUS ) CHUGAI PHARM CO LTD.

McConnell SJ, Spinella DG;

WPI; 2000-160963/14.  
N-PSDB; AA287257.

### New polypeptide useful as a receptor agonist or antagonist

wounds and promoting angiogenic capability, and as a model for designing small molecules with agonist or antagonist activity.

Claim 16; Page 65; 83pp; English

Sequences AA77154–V77174 represent novel peptides capable of binding to and activating human  $\epsilon$ -receptors.

peptides of the invention have amino acid sequences that are unrelated to that of the human FGFR or to the amino acid sequences of the other

peptides. The peptides were isolated from phage display libraries via

CC their ability to bind an FGFR probe, peptide 13-1 (AAV77161) was used as  
CC the basis for the construction of an evolved phage display library,  
CC which led to the identification of further peptides (AAV77168 -77174,  
CC consensus sequence AAV77153) capable of binding human FGFR. The peptides  
CC (or peptide mimetics) can be used to activate human FGFR, or can be used  
CC as competitive inhibitors to inhibit the binding of FGF to its receptor.  
CC The invention also encompasses methods of screening drugs that mimic  
CC human FGF. The peptides are useful as FGFR agonist or antagonist  
CC therapeutic agents for treating wounds and promoting angiogenesis. The  
CC peptides are also useful as a model for designing small molecules which  
CC have FGFR agonist or antagonist activity. The peptides are stable and  
CC economical, and are alternatives to expensive recombinant FGF or FGF  
CC isolated from animal tissue. In addition, they have an improved  
CC therapeutic delivery effect and pharmacokinetics.  
CC  
CC Sequence 15 AA;  
CC

Query Match	24.1%	Score 35;	DB 21;	Length 15;
Best Local Similarity	50.0%;	Pred. No. 33;		
Matches	7;	Conservative	2.	Worst Match

QY	Sequences	Mismatches	Indels	Gaps
14	SEISVGAFFNKDC	27		
	:       :			
2	saifvgaflhpdc	15		

```

RESULT      3
AAAE00613
ID    AAEE00613  standard, peptide; 19 AA
xx
xx

```

02-JUL-2001 (first entry)

PCBC1 APP beta-secretase cleavage site (NL/KI) mutant

Amyloid precursor protein; APP; caspase; beta-secretase; cleavage site  
 cysteine protease; apoptosis; caspase; caspase

tumour; cathepsin B; urokinase; proliferation; gene therapy; interdomain linker; Alzheimer's disease; mutant; mutation; caspase expression cassette; metastasis;

Unidentified.

Key  
Miscellaneous

```

T      /note="Wild-type Asn substituted by Lys"
T      misc-difference 11
T      misc-difference 10

```

... /note= "Wild-type Leu substituted by Ile

WO200129232-A2.

26-APR-2001.

19-OCT-2000; 2000WO-US28941.

20-OCT-1999; 99US-0160559.  
14-AUG-2000; 2000US-0225564.

(SCIO-) SCIOS INC.

Cordell B, Li Y;

WPI; 2001-290920/30.

Novel fusion polypeptide comprising first and second caspase subunit separated by cleavage site not associated in natural

subunit, useful for cloning gene encoding enzymes involved in proteolytic cleavage -

Disclosure; Fig 29A; 116pp; English.

The present amino acid sequence is a pCBCl1 plasmid amyloid precursor



CC protein (APP) beta-secretase cleavage site (AAE0611) (NL/K1) mutant. APP  
 CC sequence containing this mutation is less susceptible to beta-secretase  
 CC cleavage. APP beta-secretase cleavage site is used to construct an  
 CC artificially engineered chimeric cassette comprising human caspase-3 with  
 CC interdomain linker replaced by Swedish mutant beta-secretase cleavage  
 CC site. This modified caspase-3 plays a pivotal role in Alzheimer's  
 CC disease. Caspases are a family of cysteine proteases, that participate  
 CC in the initiation and execution of apoptosis.  
 CC The present invention relates to a method for functional cloning of genes  
 CC encoding proteins or enzymes involved in proteolytic cleavage. The  
 CC invention is based on the use of caspase expression cassettes comprising  
 CC the coding sequence of a proteolytic cleavage site flanked by sequences  
 CC encoding two caspase subunits. A fusion polypeptide comprising a first  
 CC and a second caspase subunit, separated by a cleavage site not associated  
 CC in nature, is useful for cloning gene encoding enzymes involved in  
 CC proteolytic cleavage. An expression cassette containing fusion  
 CC polypeptide is used to identify a mutant cell line deficient in an  
 CC enzyme of interest and is also useful for diagnosis and suppression of  
 CC proliferation or metastases of a tumour cell characterised by  
 CC overexpression of a polypeptide (e.g. Cathepsin B or urokinase,  
 CC selectively expressed in the tumour cells). DNA encoding fusion  
 CC polypeptide is used in gene therapy.

XX Sequence 19 AA;

Query Match 23.4%; Score 34; DB 22; Length 19;

Best Local Similarity 50.0%; Pred. No. 65;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 14 SEISVGAEFNKD 25

Db 7 sevkdaefrhd 18

RESULT 4

AAW35697 standard; peptide; 20 AA.

XX AAW35697;

XX 13-MAY-1998 (first entry)

DE D. maculata antigen 5 peptide (residues 91-110).

XX Immunomodulatory peptide; vespid antigen 5; immunogenic; allergy;

KW vespid venom; white face hornet wasp; immunodominant peptide; T cell.

XX Dolichovespula maculata.

OS Synthetic.

XX WO9733910-A1.

XX 18-SEP-1997.

XX 11-MAR-1997; 97WO-US03753.

XX 11-MAR-1996; 96US-0614935.

XX (UYRO ) UNIV ROCKEFELLER.

XX King TP.

XX MPI; 1997-470817/43.

XX Vespid venom antigen 5 peptide fragments - useful to treat or

XX diagnose vespid venom sensitivity

XX Example 1; Figure 2; 73pp; English.

XX Sequences AAW35694-99 represent 15-20 residue peptides of white faced

XX hornet antigen 5 molecule. The invention relates to peptides derived

CC from vespid venom antigen 5 (VVS) that are antigenic for T cell

CC proliferation in mice immunised with VVS. The peptides can be used to  
 CC treat or diagnose vespid venom sensitivity e.g. to Dolichovespula  
 CC maculata (white face hornet), Vespula vulgaris (yellowjacket), V.  
 CC annulifrons (yellowjacket), D. arenaria (yellow hornet), Polistes  
 CC annularis (wasp), P. exclamans (wasp), V. crabro (European hornet), V.  
 CC flavopilosa (yellowjacket), V. germanica (yellowjacket), V. pennsylvanica  
 CC (yellowjacket), V. squamosa (yellowjacket), V. vidua (yellowjacket) and  
 CC P. fuscatus (paperwasp).

XX Sequence 20 AA;

Query Match 23.4%; Score 34; DB 18; Length 20;

Best Local Similarity 71.4%; Pred. No. 70;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 21 EFNKDDC 27

Db 5 dlnhdc 11

RESULT 5

AAW82186 standard; peptide; 21 AA.

XX AAW82186;

XX 18-FEB-1999 (first entry)

DE Fluorogenic protease indicator Swedish KM/NL AMLOID peptide #1.

XX Protease activity; fluorophore; detection; fluorogenic; cellular uptake;

KW conformation change.

XX Synthetic.

XX Key Location/Qualifiers

XX Modified-site 3

XX /label= Alb

XX /note= "alpha-aminoisobutyric acid, labelled as

XX amino acid B in the specification"

XX Modified-site 4

XX /note= "epsilon-aminocaproic acid, labelled as

XX amino acid J in the specification"

XX Modified-site 16

XX /note= "epsilon-aminocaproic acid, labelled as

XX amino acid J in the specification"

XX WO9837226-A1.

XX 27-AUG-1998.

XX 20-FEB-1998; 98WO-US03000.

XX 20-FEB-1997; 97US-0802981.

XX (ONCO-) ONCOIMMUNIN INC.

XX Komoriya A, Packard BS;

XX MPI; 1998-467579/40.

XX New fluorogenic compositions - containing 2 fluorophores separated

XX by a peptide comprising a protease binding site, used for detecting

XX protease activity in samples.

XX disclosure; Page 28; 90pp; English.

XX AAW82023-W82240 are peptides used in the construction of a fluorogenic

XX composition which is used for the detection of protease activity in

XX biological samples. The products can be used for the detection of

XX conformation changes in nucleic acids, oligosaccharides,

XX polysaccharides, proteins, peptides, lipids, phospholipids, glycolipids,

CC glycoproteins, steroids or polymers. In addition, attachment of a  
 CC hydrophobic group to a molecule can be used to enhance uptake by cells.  
 CC The composition is composed of P = peptide comprising a protease binding  
 CC site for the protease, F1, F2 peptides = fluorophores where F1 is  
 CC attached to the amino terminal amino acid and F2 is attached to the  
 CC carboxyl terminal amino acid and S1, S2 peptides = when present, are  
 CC peptide spacers where S1, when present, is attached to the amino terminal  
 CC acid, and S2, when present, is attached to the carboxyl terminal amino  
 CC acid.  
 XX  
 SQ Sequence 21 AA;

Query Match 23.4%; Score 34; DB 19; Length 21;  
 Best Local Similarity 53.3%; Pred. No. 74;  
 Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1;  
 QY 14 SEISVGAEFN--KDD 26  
 ||::||| |||  
 Db 6 sevnldaefgcpkdd 20

## RESULT 6

AAG73203 ID AAG73203 standard; Peptide: 21 AA.

AC AAG73203;

DT 14-AUG-2001 (first entry)

DE Protease binding site #137.

KM Protease detection; peptide cleavage; enzyme activity; fluorogenic;  
 KW viral infection; cancer metastasis; emphysema; arthritis;  
 KM thrombosis; haemophilia.

OS Synthetic.

FH Key Location/Qualifiers  
 FT Modified-site 3 /label= Alb

FT /note= "2-aminoisobutyric acid"

PN WO200118238-A1.

PD 15-MAR-2001.

PF 11-SEP-2000; 2000WO-US24882.

PR 10-SEP-1999; 99US-0394019.

PA (ONCO-) ONCOIMMUNIN INC.

PI Komoriya A, Packard BS;

DR WPI; 2001-389573/41.

XX New fluorogenic compositions whose fluorescence level increases in the  
 PT presence of active proteases, useful for detecting and localizing  
 PT protease activity in biological samples, particularly in frozen tissue  
 PT samples -  
 XX Disclosure; Page 27; 86pp; English.

XX The present invention describes fluorogenic compositions which can be  
 CC used for the detection of protease activity. This can be useful as an  
 CC indicator of viral infection, cancer metastasis, haemophilia, emphysema,  
 CC thrombosis and arthritis. The fluorogenic compositions comprise a  
 CC peptide, a peptide spacer and a donor and an acceptor fluorophore. The  
 CC peptide is cleaved by a protease and the fluorophores can then be  
 CC detected. The present sequence is one of the peptides described in the  
 CC exemplification of the invention.

SQ Sequence 21 AA;

Query Match 23.4%; Score 34; DB 22; Length 21;  
 Best Local Similarity 53.3%; Pred. No. 74;  
 Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1;  
 QY 14 SEISVGAEFN--KDD 26  
 ||::||| |||  
 Db 6 sevnldaefgcpkdd 20

## RESULT 7

AAG73204 ID AAG73204 standard; Peptide: 21 AA.

AC AAG73204;

DT 14-AUG-2001 (first entry)

DE Protease binding site #138.

KM Protease detection; peptide cleavage; enzyme activity; fluorogenic;  
 KW viral infection; cancer metastasis; emphysema; arthritis;  
 KM thrombosis; haemophilia.

OS Synthetic.

FH Key Location/Qualifiers  
 FT Modified-site 1 /label= OTHER

FT /note= "modified by fluorophore"

FT Modified-site 3 /label= Alb

FT /note= "2-aminoisobutyric acid"

PN WO200118238-A1.

PD 15-MAR-2001.

PF 11-SEP-2000; 2000WO-US24882.

PR 10-SEP-1999; 99US-0394019.

PA (ONCO-) ONCOIMMUNIN INC.

PI Komoriya A, Packard BS;

DR WPI; 2001-389573/41.

XX New fluorogenic compositions whose fluorescence level increases in the  
 PT presence of active proteases, useful for detecting and localizing  
 PT protease activity in biological samples, particularly in frozen tissue  
 PT samples -  
 XX Disclosure; Page 27; 86pp; English.

XX The present invention describes fluorogenic compositions which can be  
 CC used for the detection of protease activity. This can be useful as an  
 CC indicator of viral infection, cancer metastasis, haemophilia, emphysema,  
 CC thrombosis and arthritis. The fluorogenic compositions comprise a  
 CC peptide, a peptide spacer and a donor and an acceptor fluorophore. The  
 CC peptide is cleaved by a protease and the fluorophores can then be  
 CC detected. The present sequence is one of the peptides described in the  
 CC exemplification of the invention.

SQ Sequence 21 AA;

Query Match 23.4%; Score 34; DB 22; Length 21;  
 Best Local Similarity 53.3%; Pred. No. 74;  
 Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 14 SEISVGAERN--KDD 26  
 11::111111  
 Db 6 sevnldaefgcpkdd 20

## RESULT 8

AA96324  
 ID AAY96324 standard; Peptide: 26 AA.

AC AAY96324;

DT 17-AUG-2000 (first entry)

DE Mammalian Bcl-2 homology domain 3 domain.

KM Mammal; apoptosis; cell death; Bcl-2; apoptosis promotion; Bcl-2

KW apoptosis inhibition; malignant cell; autoimmune disease.

OS Mammalia.

PN WO200026228-A1.

PD 11-MAY-2000.

PF 28-OCT-1999; 99WO-US25285.

PR 02-NOV-1998; 98US-0184168.

PA (CLON-) CLONTECH LAB INC.

PI Zhu L, Yin X, Chittenden T;

DR WPI: 2000-365560/31.

PT Novel polynucleotide encoding a Bcl-2 protein which is useful for  
 PT modulating apoptosis, especially in the treatment of cancer and  
 PT autoimmune diseases.

PS Disclosure: Fig 4; 47pp; English.

CC The present sequence is the mammalian Bcl-2 homology domain 3  
 CC (BH3) domain, which was used in a sequence alignment with the same  
 CC domain of a putative version of the mammalian apoptosis  
 CC regulator Bcl-2, which was designated Bcl-2-ORF2. The Bcl-2 protein,  
 CC nucleic acids and antibodies are suitable for use in promoting cell  
 CC death or for preventing apoptosis in malignant cells and those causing  
 CC autoimmune diseases.

SQ Sequence 26 AA;

Query Match 23.4%; Score 34; DB 21; Length 26;

Best Local Similarity 60.0%; Pred. No. 96;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CLGYHDLVSL 11  
 1:1:1111

Db 11 clgdemdvsl 20

## RESULT 9

AA70374  
 ID AAB70374 standard; Peptide: 26 AA.

AC AAB70374;

DT 02-MAY-2001 (first entry)

DE Bcl-2 consensus peptide sequence SEQ ID NO:7.

KM Bcl-XL/Bcl-2 associated cell death regulator; BAD; mutant; apoptosis;

KW immunostimulant; neuroprotective; nootropic; antischismatic; vulnerary;

KW cytosolic; antiviral; antitumor; antitumor; antitumor; wound healing;

KW immunosuppressive; apoptosis inducer; apoptosis inhibitor; cancer;  
 KW immunodeficiency disease; neurodegenerative disease; viral infection;  
 KW ischaemic cell death; reperfusion cell death; arthritis; infertility;  
 KW lymphoproliferative condition; inflammation; autoimmune disease.

OS Unidentified.

PN WO200110888-A1.

PD 15-FEB-2001.

PF 30-MAY-2000; 2000WO-US11864.

PR 28-MAY-1999; 99US-0136783.

PA (APOB-) APOPTOSIS TECHNOLOGY INC.

PI Zhou X;

DR WPI: 2001-138734/14.

PT New mutant Bcl-XL/Bcl-2 Associated Cell Death Regulator polypeptide,  
 PT useful for screening for candidate compounds which induce or inhibit  
 PT apoptosis, comprises amino acid substitutions at Ser118, Ser155 or  
 PT Ser113.

PS Example 2; Fig 3a; 157pp; English.

CC The present invention describes an isolated or synthetic polypeptide  
 CC (I) comprising a less than full length amino acid sequence of a mutant  
 CC Bcl-XL/Bcl-2 associated cell death regulator polypeptide (BAD) or its  
 CC fragment, which contains amino acid substitutions at Ser118 of a human  
 CC BAD, Ser155 of a murine BAD (longer murine BAD) or Ser113 of a murine  
 CC BAD (shorter murine BAD). (I) has immunostimulant, neuroprotective,  
 CC nootropic, antischismatic, vulnerary, cytosolic, antiviral,  
 CC antitumor, antitumor, antitumor, antitumor, antitumor, antitumor,  
 CC can be used as an apoptosis inducer or inhibitor. BAD polypeptides and  
 CC polynucleotides can be used for screening candidate compounds and drugs  
 CC for activity that promote cell survival or apoptosis. Other uses include  
 CC inducing or inhibiting apoptosis in a cell. Candidate compounds  
 CC identified and (mutant) BAD polypeptides are useful in treating  
 CC immunodeficiency diseases, neurodegenerative diseases, ischaemic cell  
 CC death, reperfusion cell death, wound healing, cancer, viral infections,  
 CC lymphoproliferative conditions, arthritis, infertility, inflammation and  
 CC autoimmune diseases. The present sequence represents a Bcl-family member  
 CC Bcl-2 domain consensus sequence which is used in an example from the  
 CC present invention.

SQ Sequence 26 AA;

Query Match 23.4%; Score 34; DB 22; Length 26;

Best Local Similarity 60.0%; Pred. No. 96;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CLGYHDLVSL 11  
 1:1:1111

Db 11 clgdemdvsl 20

## RESULT 10

AA37016  
 ID AAB37016 standard; peptide: 27 AA.

AC AAB37016;

DT 28-FEB-2001 (first entry)

DE Bcl2 polypeptide BH3 domain peptide #16.

KM Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;

KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;

KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;



KW	Schizophrenia; neuroleptic; diagnostic; neuropsychiatric disorder;
XX	neurological disorder; neuropathy.
OS	Homo sapiens.
PN	WO200163293-A2.
PD	30-AUG-2001.
PF	23-FEB-2001; 2001WO-GB00783.
XX	
PR	24-FEB-2000; 2000GB-0004415.
PR	28-NOV-2000; 2000US-0750395.
PA	(OXFO-) OXFORD GLYSCSCIENCES UK LTD.
PI	Herath HMAC, Parekh RB, Rohloff C;
XX	
DR	WPI: 2001-502868/55.
PT	Diagnosing and monitoring Schizophrenia by detecting the presence of Schizophrenia Associated Features and Schizophrenia Associated Protein Isoforms in samples of cerebrospinal fluid -
PS	Claim 6; Page 33; 16pp: English.
XX	
CC	The invention relates to methods and compositions for screening, diagnosis and prognosis of schizophrenia. The method involves detecting the presence of Schizophrenia (SCH) Associated Features (SFS) and SCH Associated Protein Isoforms (SPIs) in samples, e.g. by electrophoresis, immunassay or hybridisation assay, for diagnosing and monitoring SCH, studying the effectiveness of treatments and for identifying potential therapeutic agents. The method is used for (1) screening or diagnosis of SCH and the relative abundance of at least 1 chosen feature correlates with the presence or absence of SCH; and (2) monitoring the effect of therapy administered to a subject with SCH and the relative abundance of at least 1 chosen feature which correlates with the severity of SCH. The expression and activity of the SFS, SPIs and related molecules (e.g. secondary messengers) are studied to diagnose SCH, monitor the progress of the disorder and the effectiveness of treatment and as targets to identify and produce potential therapeutic agents for the treatment of SCH. The paucity of detectable neurralgic defects for distinguishes neuropsychiatric disorders such as SCH from neurological disorders, where manifestations of anatomical and biochemical changes have been identified in many cases. Consequently the identification and characterisation of cellular and/or molecular causative defects and neuropathies are necessary for improved treatment of neuropsychiatric disorders. AAU1514-AAU15762 represent the amino acid sequences of schizophrenia-associated isoforms used in the method of the invention.
CC	
XX	
SO	Sequence 13 AA:
OY	10 SLAFSEISYGAE 21
DB	1 sldfrelvvaee 12
Query Match	22.8%; Score 33; DB 22; Length 13;
Best Local Similarity	58.3%; Pred. No. 60;
Matches 7; Conservative	2; Mismatches 3; Indels 0; Gaps 0;
RESULT 13	
AAW38432	
ID AAW38432 standard; Peptide: 15 AA.	
XX AAW38432;	
DT 21-MAY-1998 (first entry)	
DE Human dendritic cell receptor DEC-205 peptide 3 (aa82-96).	
XX Dendritic cell receptor; DEC-205; human; ligand; cell targeting;	

KM	antigen: toxin; immunosuppressive; therapy; prophylaxis;
KM	monoclonal antibody.
XX	
OS	Synthetic.
OS	Homo sapiens.
XX	
PN	W09745449-A1..
XX	
PD	04-DEC-1997.
XX	
XX	29-MAY-1997; 97WO-N200068.
PF	
XX	
PR	29-MAY-1996; 96NZ-028692.
XX	
PA	(HART/) HART D N J.
XX	
PI	Hart DNJ;
XX	
DR	WPI: 1998-032580/03.
XX	
PT	Human dendritic cell receptor, DEC-205 - useful to develop ligands
PR	for isolation and targeted cell delivery of antigen or toxin
XX	
PS	Example 3; Fig 11; 53pp; English.
XX	
CC	DEC-205 peptide 3 corresponds to amino acids 82-96 of the human
CC	dendritic cell receptor DEC-205 (see AAW38429). BALB/c mice have
CC	been immunised with DEC-205 peptide 1 (see AAW38430) and DEC-205
CC	peptide 2 (see AAW38431). Spleen cells were subsequently fused with
CC	NS-1 myeloma cells. A hybridoma, 2F5, was obtained with produced
CC	monoclonal antibody binding to peptide 1 but not to peptide 2 or
CC	to control DEC-205 peptide 3. The invention provides isolated
CC	human DEC-205, its extracellular domain and equivalent fragments,
CC	and polynucleotides encoding these polypeptides. Further provided
CC	are ligands (preferably an antibody or antibody binding fragment)
CC	that bind to human DEC-205, as well as constructs for use in
CC	prophylaxis and therapy that comprise such a ligand, human DEC-205
CC	or an extracellular domain coupled to an antigen capable of
CC	inducing a protective immune response in a patient, or to a toxin
CC	(e.g. ricin A chain to specifically destroy dendritic cells as part
CC	of an immunosuppressive process).
XX	
SQ	Sequence 15 AA;
Query Match	22.8%; Score 33; DB 19; Length 15;
Best Local Similarity	43.8%; Pred. No. 72;
Matches 7; Conservativity 5; Mismatches 2; Indels 2; Gaps 1;	
OY	1 KCLGYHLDVSLAFSEI 16
	:::   :
DB	1 Kclg--ldltkynel 14
RESULT 14	
AAB06316	
ID	AAB06316 standard; peptide; 16 AA.
XX	
AC	AAB06316;
XX	
DT	03-OCT-2000 (first entry)
XX	
DE	Human beta-amyloid precursor protein beta-secretase site.
XX	
KM	Human; beta-amyloid precursor protein; beta-APP; beta-secretase;
KM	subtilisin-kevin isoenzyme 1; SKI-1;
KM	pro-brain-derived neurotrophic factor; PROBDNF; antiipaeic;
KM	cytostatic; vasotropic; SKI-1 inhibitor; hypercholesterolaemia;
KM	liver steatosis; Ras-dependent cancer; restenosis;
KM	amyloid protein formation.
XX	
OS	Mammalia.
XX	

XX (SCIO-) SCIOS INC.  
PA  
XX  
PI Cordell B, Li Y;  
PI  
DR WPI: 2001-290920/30.  
XX  
PT Novel fusion polypeptide comprising first and second caspase subunit  
PT separated by cleavage site not associated in nature with caspase  
PT subunit, useful for cloning gene encoding enzymes involved in  
PT proteolytic cleavage -  
XX  
XX  
XX Example 2; Page 26; 116pp; English.  
XX  
XX The present sequence is a beta-secretase cleavage site of beta-amyloid  
CC precursor protein (beta-Ap) mutant found in certain families of  
CC autosomal dominant form of Alzheimer's disease. This sequence is used to  
CC construct a chimeric cassette comprising human caspase-3 with interdomainal  
CC linker replaced by this sequence. This modified caspase-3 plays a pivotal  
CC role in Alzheimer's disease. Caspases are a family of cysteine proteases  
CC that participate in the initiation and execution of apoptosis.  
CC The present invention relates to a method for functional cloning of genes  
CC encoding proteins or enzymes involved in proteolytic cleavage. The  
CC invention is based on the use of caspase expression cassettes comprising  
CC the coding sequence of a proteolytic cleavage site flanked by sequences  
CC encoding two caspase subunits. A fusion polypeptide comprising a first  
CC and a second caspase subunit, separated by a cleavage site not associated  
CC in nature, is useful for cloning gene encoding enzymes involved in  
CC proteolytic cleavage. An expression cassette containing fusion  
CC polypeptide is used to identify a mutant cell line deficient in an  
CC enzyme of interest and is also useful for diagnosis and suppression of  
CC overexpression or metastases of a tumour cell characterised by  
CC overexpression of a polypeptide (e.g. Cathepsin B or tyrosinase,  
CC selectively expressed in the tumour cells). DNA encoding fusion  
XX polypeptide is used in gene therapy.  
XX

```

Query Match          22.88;   Score 33;   DB 22;   Length 18;
Best Local Similarity 50.0%;
Matches      6;   Conservative      3;   Mismatches      3;   Indels      0;   Gaps      0;

OY      14 SEISVGAEFNKD 25
      ||::||| |
Db      6 sevnldaeifnd 17

RESULT 16
AAE00611
ID   AAE00611 standard; peptide; 19 AA.
XX
AC   AAE00611;
XX
DT      02-JUL-2001 (first entry)
XX
DE   Amyloid precursor protein (APP) beta-secretase cleavage site.
XX
KW   Amyloid precursor protein; APP; caspase; beta-secretase; cleavage site;
KW   cysteine protease; apoptosis; caspase expression cassette; metastasis;
KW   tumour; cathepsin B; urokinase; proliferation; gene therapy;
KW   interdomain linker; Alzheimer's disease.
XX
OS   unidentified.
XX
PN   WO200129232-A2.
XX
PD      26-APR-2001.
XX
PF      19-OCT-2000; 2000WO-US28941.
XX
PR      20-OCT-1999; 99US-0160559.
PR      14-AUG-2000; 2000US-0225564.
PR

```

```

XX (SCIO-) SCIOS INC.
PA
XX
PI Cordell B, Li Y;
XX
DR WPI: 2001-290920/30.
XX
XX Novel fusion polypeptide comprising first and second caspase subunit
PT separated by cleavage site not associated in nature with caspase
PT subunit, useful for cloning gene encoding enzymes involved in
PT proteolytic cleavage
XX
PS Disclosure; Fig 28A; 116pp; English.
XX
XX The present amino acid sequence is a beta-secretase cleavage site of
CC an amyloid precursor protein (APP). This sequence is used to construct
CC an artificially engineered chimeric cassette comprising human caspase-3
CC with interdomain linker replaced by Swedish mutant beta-secretase
CC cleavage site. This modified caspase-3 plays a pivotal role in
CC Alzheimer's disease. Caspases are a family of cysteine proteases, that
CC participate in the initiation and execution of apoptosis.
CC The present invention relates to a method for functional cloning of genes
CC encoding proteins or enzymes involved in proteolytic cleavage. The
CC invention is based on the use of caspase expression cassettes comprising
CC the coding sequence of a proteolytic cleavage site flanked by sequences
CC encoding two caspase subunits. A fusion polypeptide comprising a first
CC and a second caspase subunit, separated by a cleavage site not associated
CC in nature, is useful for cloning gene encoding enzymes involved in
CC proteolytic cleavage. An expression cassette containing fusion
CC polypeptide is used to identify a mutant cell line deficient in an
CC enzyme of interest and is also useful for diagnosis and suppression of
CC proliferation or metastases of a tumour cell characterised by
CC overexpression of a polypeptide (e.g. Cathepsin B or urokinase,
CC selectively expressed in the tumour cells). DNA encoding fusion
CC polypeptide is used in gene therapy.
XX
SQ Sequence 19 AA:

Query Match 22.8%; Score 33; DB 22; Length 19;
Best Local Similarity 50.0%; Pred. No. 96;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 14 SEISVGAEFNKD 25
   ||::||| |
Db 7 sevnldaefrhd 18

RESULT 17
AAy69714
ID AAY69714 standard; peptide; 20 AA.
XX
AC AAY69714;
XX
DT 11-APR-2000 (first entry)
XX
DE Beta-APP alpha-secretase substrate [NL]-APP(-10,+10).
XX
KM Neotropic; neuroprotective; beta-amyloid precursor protein; metabolism;
KM cleavage site; beta-secretase; neurodegenerative disease;
KM Alzheimer's disease.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX WO9964587-A1.
XX
XX 16-DEC-1999.
XX
XX 04-JUN-1999; 99WO-FR01326.
XX
XX 05-JUN-1998; 98FR-0007068.
PR 31-MAR-1999; 99US-0122599.
PR

```

```

XX (RHON ) RHONE-POULENC RORER SA.
PA
XX (UYPA-) UNIV CURIE PARIS VI P & M.
XX
PI Rholam M, Munoz-Gimenez N, Moutaouakil M, Cohen P, Bertrand P;
XX
DR WPI: 2000-097537/08.
XX
XX Polypeptide with beta-secretase activity, specific for wild-type
PT amyloid precursor protein, useful in treating Alzheimer's disease
PT
XX
XX Example 3; Page 24; 44pp; French.
XX
XX Peptides AAY69702-Y69718 represent synthetic peptide substrates for a
CC novel polypeptide with beta-secretase activity that can cleave
CC specifically the natural beta-amyloid precursor protein (bAPP). Normal
CC cleavage of the protein occurs between amino acids Met596-Asp597 and
CC Val636-Ile637 (positions 4-5 and 44-45 of AAY69701). The novel
CC polypeptide is used to identify agents that interact specifically with
CC it. These agents regulate metabolism of APP, particularly they slow down
CC or reduce production of beta-amyloid, so can be used to treat
CC neurodegenerative diseases, particularly Alzheimer's disease.
XX
SQ Sequence 20 AA:

Query Match 22.8%; Score 33; DB 21; Length 20;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 14 SEISVGAEFNKD 25
   ||::||| |
Db 6 sevnldaefrhd 17

RESULT 18
AAy69716
ID AAY69716 standard; peptide; 20 AA.
XX
AC AAY69716;
XX
DT 11-APR-2000 (first entry)
XX
DE Beta-APP alpha-secretase substrate [NM]-APP(-10,+10).
XX
KM Neotropic; neuroprotective; beta-amyloid precursor protein; metabolism;
KM cleavage site; beta-secretase; neurodegenerative disease;
KM Alzheimer's disease.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX WO9964587-A1.
XX
XX 16-DEC-1999.
XX
XX 04-JUN-1999; 99WO-FR01326.
XX
XX 05-JUN-1998; 98FR-0007068.
PR 31-MAR-1999; 99US-0122599.
XX
XX (RHON ) RHONE-POULENC RORER SA.
PA (UYPA-) UNIV CURIE PARIS VI P & M.
XX
XX Rholam M, Munoz-Gimenez N, Moutaouakil M, Cohen P, Bertrand P;
XX
DR WPI: 2000-097537/08.
XX
XX Polypeptide with beta-secretase activity, specific for wild-type
PT amyloid precursor protein, useful in treating Alzheimer's disease
PT
XX Example 3; Page 24; 44pp; French.
XX

```

CC Peptides AAY69702-Y69718 represent synthetic peptide substrates for a  
 CC novel polypeptide with beta-secretase activity that can cleave  
 CC specifically the natural beta-amyloid precursor protein (bAPP). Normal  
 CC cleavage of the protein occurs between amino acids Met596-Asp597 and  
 CC Val636-Ile637 (positions 4-5 and 44-45 of AAY69701). The novel  
 CC polypeptide is used to identify agents that interact specifically with  
 CC it. These agents regulate metabolism of APP, particularly they slow down  
 CC or reduce production of beta-amyloid, so can be used to treat  
 CC neurodegenerative diseases, particularly Alzheimer's disease.  
 XX Sequence 20 AA;

Query Match 22.8%; Score 33; DB 21; Length 20;  
 Best Local Similarity 50.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 14 SEISVGAEPNKD 25  
 ||::||| |  
 Db 6 sevnmdaefrhd 17

## RESULT 19

AAW08361  
 ID AAW08361 standard; peptide; 21 AA.

AC AAW08361;

DT 05-SEP-1997 (first entry)

DE Beta-secretase substrate #3.

KW Beta-cleavage site; beta amyloid precursor protein; APP; beta-secretase;  
 KW alpha-secretase; proteolytic cleavage; inhibitor; Alzheimer's disease.

OS Synthetic.

XX PN W09640885-A2.

XX PD 19-DEC-1996.

XX PF 07-JUN-1996; 96WO-US09985.

XX PR 07-JUN-1995; 95US-0485152.

XX PR 07-JUN-1995; 95US-0480498.

XX PA (ATHE-) ATHENA NEUROSCIENCES INC.

XX PI Anderson JP, Chrysler SMS, Jacobson-croak KL, Kelm PS;  
 XX PI Mcconlogue LC, Sinha S, Tan H;

XX DR WPI; 1997-052304/05.

XX PT Beta-secretase which specifically cleaves beta-amyloid precursor  
 XX PT protein - useful to screen for inhibitors useful in treatment of  
 XX PT Alzheimer's disease

XX PS Disclosure; Page 45; 92pp; English.

XX CC AAW08359-W08362 represent substrates for the enzyme of the invention.  
 CC The enzyme of the invention is beta-secretase, and specifically cleaves  
 CC beta-amyloid precursor protein (beta-APP). Normal processing of beta-APP  
 CC is thought to occur via cleavage between residues 16 and 17 of the  
 CC beta-amyloid peptide region by an alpha-secretase. Pathogenic processing  
 CC is thought to occur by beta-secretase cleavage of beta-APP.  
 CC Beta-secretase activity can be detected and measured using a method of  
 CC the invention, which detects at least one of the beta-secretase cleavage  
 CC products formed on cleavage. The method can be used to determine whether  
 CC a test substance inhibits proteolytic cleavage, by beta-secretase, of  
 CC beta-APP. Compounds inhibits proteolytic cleavage, by beta-secretase, of  
 CC beta-secretase activity can be used to inhibit cleavage of beta-APP in  
 CC cells or mammalian hosts. Isolation and purification of beta-secretase  
 CC will permit chemical modelling of a critical event in the pathology of

CC Alzheimer's disease.  
 XX Sequence 21 AA;  
 SO

Query Match 22.8%; Score 33; DB 18; Length 21;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 14 SEISVGAEPNKD 25  
 ||::||| |  
 Db 1 sevnldaeifhd 12

## RESULT 20

AAW82188  
 ID AAW82188 standard; peptide; 21 AA.

AC AAW82188;

DT 18-FEB-1999 (first entry)

DE Fluorogenic protease indicator native amyloid peptide.

KW Protease activity; fluorophore; detection; fluorogenic; cellular uptake;  
 KW conformation change.

OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Modified-site 3

XX FT /label= Aib

XX FT /note= "alpha-aminoisobutyric acid, labelled as

XX FT Modified-site 4

XX FT /note= "epsilon-aminocaproic acid, labelled as

XX FT Modified-site 16

XX FT /note= "epsilon-aminocaproic acid, labelled as

XX FT amino acid J in the specification"

XX PN W09837226-A1.

XX PD 27-AUG-1998.

XX PF 20-FEB-1998; 98WO-US03000.

XX PR 20-FEB-1997; 97US-0802981.

XX PA (ONCO-) ONCOIMMUNIN INC.

XX PI Komoriya A, Packard BS;

XX DR WPI; 1998-467579/40.

XX PT New fluorogenic compositions - containing 2 fluorophores separated  
 XX PT by a peptide comprising a protease binding site, used for detecting  
 XX PT protease activity in samples.

XX PS Disclosure; Page 29; 90pp; English.

XX CC AAW82023-W82240 are peptides used in the construction of a fluorogenic  
 CC composition which is used for the detection of protease activity in  
 CC biological samples. The products can be used for the detection of  
 CC conformational changes in nucleic acids, oligosaccharides,  
 CC polysaccharides, proteins, peptides, lipids, phospholipids, glycolipids,  
 CC glycoproteins, steroids or polymers. In addition, attachment of a  
 CC hydrophobic group to a molecule can be used to enhance uptake by cells.  
 CC The composition is composed of P = peptide comprising a protease binding  
 CC site for the protease, F1, F2 peptides = fluorophores where F1 is  
 CC attached to the amino terminal amino acid and F2 is attached to the  
 CC carboxyl terminal amino acid and S1, S2 peptides = when present, are  
 CC peptide spacers where S1, when present, is attached to the amino terminal



CC acid, and S2, when present, is attached to the carboxyl terminal amino  
 CC acid.  
 XX  
 SQ Sequence 21 AA;

Query Match 22.8%; Score 33; DB 19; Length 21;  
 Best Local Similarity 53.3%; Pred. No. 1.1e+02;  
 Matches 8; Conservative 2; Mismatches 3; Indels 2; Gaps 1;  
 QY 14 SEISVGAEFN--KDD 26  
 ||: : ||| |||  
 Db 6 sevkmdaefgpxkdd 20

RESULT 21  
 AAY33755  
 ID AAY33755 standard; Protein; 21 AA.  
 XX  
 AC AAY33755;  
 XX  
 DF 09-NOV-1999 (first entry)  
 XX  
 DE Synthetic oligopeptide 5-16'SW.  
 XX  
 KM Beta-secretase; beta-amyloid protein precursor; APP; Down's syndrome;  
 KM Alzheimer's disease; measure activity; cleavage site.  
 XX  
 OS Synthetic.  
 XX  
 PN US5942400-A.  
 XX  
 PD 24-AUG-1999.  
 XX  
 PE 07-JUN-1996; 96US-0659984.  
 XX  
 PR 07-JUN-1996; 96US-0659984.  
 PR 07-JUN-1995; 95US-0480498.  
 PR 07-JUN-1995; 95US-0485152.  
 XX  
 PA (ELAN-) ELAN PHARM INC.  
 PI  
 PI Anderson JP, Jacobson-Croak KL, Sinha S;  
 DR WPI; 1999-517417/43.  
 XX  
 PT A method for detecting human beta-secretase cleavage of polypeptides  
 PT useful for identifying beta-secretase inhibitors  
 XX  
 PS Examples; Column 30; 43pp; English.  
 XX  
 CC Sequences AAY33752-Y33756 are synthetic oligopeptides used for measuring  
 CC the activity of beta-secretase (AAY33741). Beta-secretase is capable of  
 CC cleaving beta-amyloid protein precursor (APP) (AAY33742). These  
 CC synthetic peptides contain the cleavage site of APP. Beta-secretase and  
 CC APP are used in a method for detecting human beta-secretase cleavage of  
 CC polypeptides and for identifying beta-secretase inhibitors. Inhibition  
 CC of beta-secretase activity would be useful for chemical modelling of a  
 CC critical event in the pathology of Alzheimer's disease. Inhibitors of  
 CC beta-secretase would be useful for the prevention and treatment of  
 CC Alzheimer's disease and Down's Syndrome.  
 CC  
 SQ Sequence 21 AA;

Query Match 22.8%; Score 33; DB 20; Length 21;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 14 SEISVGAEFNKD 25  
 ||: : ||| |  
 Db 1 sevnldaefrhd 12

RESULT 22  
 AAG73205  
 ID AAG73205 standard; Peptide; 21 AA.  
 XX

AC AAG73205;  
 XX  
 DT 14-AUG-2001 (first entry)  
 XX  
 DE Protease binding site #139.  
 XX  
 KM Protease detection; peptide cleavage; enzyme activity; fluorogenic;  
 KM viral infection; cancer metastasis; emphysema; arthritis;  
 KM thrombosis; haemophilia.  
 XX

OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 3 /label= Alb  
 FT FT /note= "2-aminoisobutyric acid"  
 XX

PN MO200118238-A1.  
 XX  
 PD 15-MAR-2001.  
 XX  
 PF 11-SEP-2000; 2000WO-US24882.  
 XX  
 PR 10-SEP-1999; 99US-0394019.  
 XX

PA (ONCO-) ONCOIMMUNIN INC.  
 XX  
 PI Komoriya A, Packard BS;  
 PI  
 DR WPI; 2001-389573/41.  
 XX

XX New fluorogenic compositions whose fluorescence level increases in the  
 PT presence of active proteases, useful for detecting and localizing  
 PT protease activity in biological samples, particularly in frozen tissue  
 PT samples  
 XX  
 PS Disclosure; Page 27; 86pp; English.

XX The present invention describes fluorogenic compositions which can be  
 CC used for the detection of protease activity. This can be useful as an  
 CC indicator of viral infection, cancer metastasis, haemophilia, emphysema,  
 CC thrombosis and arthritis. The fluorogenic compositions comprise a  
 CC peptide, a peptide spacer and a donor and an acceptor fluorophore. The  
 CC peptide is cleaved by a protease and the fluorophores can then be  
 CC detected. The present sequence is one of the peptides described in the  
 CC exemplification of the invention.  
 CC  
 SQ Sequence 21 AA;

Query Match 22.8%; Score 33; DB 22; Length 21;  
 Best Local Similarity 53.3%; Pred. No. 1.1e+02;  
 Matches 8; Conservative 2; Mismatches 3; Indels 2; Gaps 1;  
 QY 14 SEISVGAEFN--KDD 26  
 ||: : ||| |||  
 Db 6 sevkldaefgpxkdd 20

RESULT 23  
 AAG73206  
 ID AAG73206 standard; Peptide; 21 AA.  
 XX  
 AC AAG73206;  
 XX  
 DT 14-AUG-2001 (first entry)  
 XX  
 DE Protease binding site #140.

XX Protease detection; peptide cleavage; enzyme activity; fluorogenic;  
 KW viral infection; cancer metastasis; emphysema; arthritis;  
 KW thrombosis; haemophilia.  
 OS Synthetic.  
 XX  
 XX  
 FT Key Location/Qualifiers  
 FT Modified-site 3 /label= Aib  
 FT /note="2-aminoisobutyric acid"  
 PN WC200118238-A1.  
 XX  
 PD 15-MAR-2001.  
 XX  
 PF 11-SEP-2000; 2000WO-US24882.  
 XX  
 PR 10-SEP-1999; 99US-0394019.  
 XX  
 PA (ONCO-) ONCOIMMUNIN INC.  
 XX  
 PI Komoriya A, Packard BS;  
 XX  
 DR WPI; 2001-389573/41.  
 XX  
 PT New fluorogenic compositions whose fluorescence level increases in the  
 PT presence of active proteases, useful for detecting and localizing  
 PT protease activity in biological samples, particularly in frozen tissue  
 PT samples -  
 XX  
 PS Disclosure; Page 27; 86pp; English.  
 XX  
 CC The present invention describes fluorogenic compositions which can be  
 CC used for the detection of protease activity. This can be useful as an  
 CC indicator of viral infection, cancer metastasis, haemophilia, emphysema,  
 CC thrombosis and arthritis. The fluorogenic compositions comprise a  
 CC peptide, a peptide spacer and a donor and an acceptor fluorophore. The  
 CC peptide is cleaved by a protease and the fluorophores can then be  
 CC detected. The present sequence is one of the peptides described in the  
 CC exemplification of the invention.  
 XX  
 SQ Sequence 21 AA;  
 SQ

Query Match 22.8%; Score 33; DB 22; Length 21;  
 Best Local Similarity 53.3%; Pred. No. 1.1e+02;  
 Matches 8; Conservative 2; Mismatches 3; Indels 2; Gaps 1;  
 QY 14 SEISVGAFFN--KDD 26  
 ||: ||| |||  
 Db 6 sevkmdaefgcpkd 20

RESULT 24  
 AAG73316  
 ID AAG73316 standard; Peptide; 21 AA.  
 XX  
 AC AAG73316;  
 XX  
 DT 14-AUG-2001 (first entry)  
 XX  
 DE Protease indicator compound peptide #45.  
 XX  
 KW Protease detection; peptide cleavage; enzyme activity; fluorogenic;  
 KW viral infection; cancer metastasis; emphysema; arthritis;  
 KW thrombosis; haemophilia.  
 XX  
 OS Synthetic.  
 XX  
 FT Key Location/Qualifiers  
 FT Modified-site 1 /label= OTHER

FT Modified-site /note="modified by fm"  
 FT 4  
 FT /label= OTHER  
 FT /note="designated J in the specification"  
 FT Misc-difference 7 /note= "D-form residue"  
 FT 10  
 FT Misc-difference 10 /note= "D-form residue"  
 FT 14  
 FT Misc-difference 14 /note= "D-form residue"  
 FT 16  
 FT Modified-site /label= OTHER  
 FT /note="designated C5 in the specification"  
 PN WC200118238-A1.  
 XX  
 PD 15-MAR-2001.  
 XX  
 PF 11-SEP-2000; 2000WO-US24882.  
 XX  
 PR 10-SEP-1999; 99US-0394019.  
 XX  
 PA (ONCO-) ONCOIMMUNIN INC.  
 XX  
 PI Komoriya A, Packard BS;  
 XX  
 DR WPI; 2001-389573/41.  
 XX  
 PT New fluorogenic compositions whose fluorescence level increases in the  
 PT presence of active proteases, useful for detecting and localizing  
 PT protease activity in biological samples, particularly in frozen tissue  
 PT samples -  
 XX  
 PS Claim 4; Page 71; 86pp; English.  
 XX  
 CC The present invention describes fluorogenic compositions which can be  
 CC used for the detection of protease activity. This can be useful as an  
 CC indicator of viral infection, cancer metastasis, haemophilia, emphysema,  
 CC thrombosis and arthritis. The fluorogenic compositions comprise a  
 CC peptide, a peptide spacer and a donor and an acceptor fluorophore. The  
 CC peptide is cleaved by a protease and the fluorophores can then be  
 CC detected. The present sequence is one of the peptides described in the  
 CC exemplification of the invention.  
 XX  
 SQ Sequence 21 AA;  
 SQ

Query Match 22.8%; Score 33; DB 22; Length 21;  
 Best Local Similarity 53.3%; Pred. No. 1.1e+02;  
 Matches 8; Conservative 2; Mismatches 3; Indels 2; Gaps 1;  
 QY 14 SEISVGAFFN--KDD 26  
 ||: ||| |||  
 Db 6 sevkmdaefgcpkd 20

RESULT 25  
 AAB47265  
 ID AAB47265 standard; Peptide; 21 AA.  
 XX  
 AC AAB47265;  
 XX  
 DT 18-JUL-2001 (first entry)  
 XX  
 DE Peptide 5-16'SW, for used in beta-secretase assay.  
 XX  
 KW Beta-secretase; isotype; beta-amyloid precursor protein; app;  
 KW beta-amyloid peptide; beta-AP; Alzheimer's disease; Downs syndrome;  
 KW HCMA-D; Swedish mutation; maltose binding protein; MBP.  
 XX  
 OS Synthetic.  
 XX  
 PN US6221645-B1.

```

XX 24-APR-2001.
PD
XX
XX 07-JUN-1996; 96US-0660531.
PF
XX 07-JUN-1995; 95US-0480498.
PR
XX
XX (ELAN-) ELAN PHARM INC.
PA
XX
XX Chrysler SMS, Sinha S, Keim PS, Anderson JP, Tan H, McConlogue LC;
XX
XX WPI; 2001-315578/33.
DR
XX
XX Novel antibody that specifically binds native beta-secretase protein,
PR useful for raising anti-idiotypic antibodies and for detecting or
PT diagnosing pathological conditions related to presence of respective
PT antigens
XX
XX Example: Column 30; 42pp; English.
PS
XX
XX The sequences given in ABA47262-67 represent synthetic peptides
CC containing the cleavage sites derived from wild-type beta-amyloid
CC precursor protein (APP). These peptides were used in assays utilizing
CC partially purified beta-secretase to identify beta-secretase inhibitors.
CC Beta-secretase is thought to be responsible for the pathogenic
CC processing of APP to form beta amyloid peptide (beta-AP) in beta-AP
CC related conditions, e.g. Alzheimer's disease, Downs syndrome, HCHWA-D
CC etc. Beta-secretase has a molecular weight of 260-300 KD and will bind
CC to wheat germ agglutinin but not to concanavalin A. Beta-secretase
CC will cleave both the wild type and the Swedish mutation of APP.
XX
XX SQ Sequence 21 AA;

```

```

Query Match 22.8%; Score 33; DB 22; Length 21;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
OY 14 SEISVGAEFNKD 25
| | : : | | | |
Db 1 sevnldaefrhd 12

```

Search completed: January 6, 2002, 09:12:21  
 Job time: 434 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:06:02 ; Search time 16.63 Seconds  
(without alignments)  
36.536 Million cell updates/sec

Title: US-09-020-393b-18

Perfect score: 145

Sequence: 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 127597

Minimum DB seq length: 0

Maximum DB seq length: 27

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfillsl.pep:\*

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	26.2	16	5 PCT-US94-07043A-1	Sequence 1, Appl
2	34	23.4	20	1 US-08-614-935-17	Sequence 17, Appl
3	34	23.4	20	1 US-09-130-287-17	Sequence 17, Appl
4	34	23.4	20	4 US-09-236-385A-41	Sequence 41, Appl
5	34	23.4	21	3 US-08-802-981-112	Sequence 112, Appl
6	33	22.8	21	2 US-08-659-984A-18	Sequence 18, Appl
7	33	22.8	21	3 US-08-802-981-113	Sequence 113, Appl
8	33	22.8	21	3 US-08-802-981-114	Sequence 114, Appl
9	33	22.8	21	4 US-08-660-531-18	Sequence 18, Appl
10	32	22.1	22	2 US-08-313-200-8	Sequence 8, Appl
11	32	22.1	27	1 US-08-141-324-11	Sequence 11, Appl
12	32	22.1	27	1 US-08-141-324-12	Sequence 12, Appl
13	32	22.1	27	1 US-08-541-902-11	Sequence 11, Appl
14	32	22.1	27	1 US-08-541-902-12	Sequence 12, Appl
15	31	21.4	19	1 US-08-324-301-3	Sequence 3, Appl
16	30	20.7	13	2 US-08-706-741B-44	Sequence 44, Appl
17	30	20.7	13	2 US-08-924-695A-44	Sequence 44, Appl
18	30	20.7	15	1 US-08-440-391-7	Sequence 7, Appl
19	30	20.7	15	1 US-08-440-391-32	Sequence 32, Appl
20	30	20.7	15	2 US-08-248-839C-152	Sequence 152, Appl
21	30	20.7	15	2 US-08-808-597A-7	Sequence 7, Appl
22	30	20.7	15	2 US-08-808-597A-32	Sequence 32, Appl
23	30	20.7	15	4 US-09-236-385A-7	Sequence 7, Appl
24	30	20.7	15	4 US-09-236-385A-32	Sequence 32, Appl
25	30	20.7	15	5 PCT-US96-06122-7	Sequence 7, Appl
26	30	20.7	15	5 PCT-US96-06122-32	Sequence 32, Appl
27	30	20.7	18	5 PCT-US94-01234-21	Sequence 21, Appl

28	29	20.0	16	4 US-09-171-705-41	Sequence 41, Appl
29	29	20.0	24	1 US-08-461-597-4	Sequence 4, Appl
30	29	20.0	24	2 US-08-535-298-4	Sequence 4, Appl
31	29	20.0	24	5 PCT-US94-05569A-4	Sequence 4, Appl
32	29	20.0	24	5 PCT-US94-05569A-4	Sequence 4, Appl
33	29	20.0	26	1 US-08-475-989-25	Sequence 25, Appl
34	29	20.0	26	2 US-08-475-989-25	Sequence 25, Appl
35	29	20.0	26	3 US-08-256-839-25	Sequence 25, Appl
36	28	19.3	9	3 US-08-802-981-219	Sequence 19, Appl
37	28	19.3	10	2 US-08-659-984A-19	Sequence 19, Appl
38	28	19.3	10	4 US-08-660-531-19	Sequence 19, Appl
39	28	19.3	11	5 PCT-US94-07043A-3	Sequence 3, Appl
40	28	19.3	19	1 US-07-678-974D-67	Sequence 67, Appl
41	28	19.3	19	2 US-08-985-090-8	Sequence 8, Appl
42	28	19.3	19	3 US-09-165-543-8	Sequence 8, Appl
43	28	19.3	19	4 US-09-165-543-15	Sequence 15, Appl
44	28	19.3	20	4 US-09-007-905-59	Sequence 59, Appl
45	28	19.3	20	4 US-09-007-905-64	Sequence 64, Appl

#### ALIGNMENTS

RESULT 1  
PCT-US94-07043A-1  
Sequence 1, Application PC/TUS9407043A  
GENERAL INFORMATION:  
APPLICANT: Tamburini, Paul P.; Benz, Ginter; H. Bich,  
Dieter; Dreyer, Robert N.; Koenig, Gerhard  
TITLE OF INVENTION: CATHEPSIN D IS AN AMYLOIDGENIC  
PROTEASE IN ALZHEIMER S DISEASE  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Miles Inc.  
STREET: 400 Morgan Lane  
CITY: West Haven  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06516  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 800 kb storage  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/07043A  
FILING DATE: June 21, 1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/10889  
FILING DATE: November 12, 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/995,660  
FILING DATE: December 16, 1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/880,914  
FILING DATE: May 11, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Pamela A. Simonon  
REGISTRATION NUMBER: 31,060  
REFERENCE/DOCKET NUMBER: MTI 224.3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 937-2795  
TELEFAX: (203) 937-2795  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
PCT-US94-07043A-1

Query Match 26.2% Score 38; DB 5; Length 16;

Best Local Similarity 53.8%; Pred. No. 4.1;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 14 SEISGAEFNKDD 26  
Db 2 SEVKDAEFRRHD 14

## RESULT 2

US-08-614-935-17  
; Sequence 17, Application US/08614935  
; Patent No. 5804201  
; GENERAL INFORMATION:  
; APPLICANT: King, Te P.  
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES OF VESPID  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; CITY: Floor  
; STATE: Hackensack  
; COUNTRY: New Jersey  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/614,935  
; FILING DATE: 11-MAR-1996  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-156  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: Internal  
; US-08-614-935-17

Query Match 23.4%; Score 34; DB 1; Length 20;  
Best Local Similarity 71.4%; Pred. No. 25;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 21 EFNKDDC 27  
Db 5 DFNHDC 11

## RESULT 3

US-09-130-287-17  
; Sequence 17, Application US/09130287  
; Patent No. 6106844  
; GENERAL INFORMATION:  
; APPLICANT: King, Te P.  
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES OF VESPID  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; CITY: Floor  
; STATE: Hackensack  
; COUNTRY: New Jersey  
; ZIP: 07601

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/130,287  
; FILING DATE:

CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/614,935  
; FILING DATE: 11-MAR-1996

CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-156 DIV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684

INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: Internal

US-09-130-287-17

Query Match 23.4%; Score 34; DB 3; Length 20;  
Best Local Similarity 71.4%; Pred. No. 25;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 21 EFNKDDC 27  
Db 5 DFNHDC 11

## RESULT 4

US-09-236-385A-41  
; Sequence 41, Application US/09236385A  
; Patent No. 6221615  
; GENERAL INFORMATION:  
; APPLICANT: CHITTENDEN, Thomas D.; and  
; LUTZ, Robert J.  
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
; MODULATE APOPTOSIS  
; NUMBER OF SEQUENCES: 41  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hale and Dorr  
; STREET: 1455 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/236,385A  
; FILING DATE: 25-Jan-1999

```

CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: WIXON, HENRY N.
REGISTRATION NUMBER: 32,073
(C) ATTORNEY DOCKET NO. 104322.147CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-942-8400
TELEFAX: 202-942-8484
INFORMATION FOR SEQ ID NO: 41
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 41
US-09-236-385A-41

Query Match      23.4%   Score 34;   DB 4;   Length 20;
Best Local Similarity 60.0%;   Pred. No. 25;
Matches      6;   Conservative      2;   Mismatches      2;   Indels      0;   Gaps      0;

QY      2  CLEGHLDVSL 11
      1 1 :11111
Db      11  CIGDEMDVSL 20

RESULT      5
US-08-802-981-112
Sequence 112: Application US/08802981
Patent No. 6037137
GENERAL INFORMATION:
APPLICANT: Komoriya, Akira
APPLICANT: Packard, Beverly S.
TITLE OF INVENTION: Compositions for the Detection of Enzyme
TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
NUMBER OF SEQUENCES: 231
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/802,981
FILING DATE: 20-FEB-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 016865-000300US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 3
OTHER INFORMATION: /product= "Aib"
FEATURE:
NAME/KEY: Modified-site

```

```

: LOCATION: 4 /product= "Acp"
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 16
: OTHER INFORMATION: /product= "Acp"
US-08-802-981-112

Query Match 23.4%; Score 34; DB 3; Length 21;
Best Local Similarity 53.3%; Pred. No. 27;
Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 14 SEISYGAEFN-KDD 26
|::: ||| |||
Db 6 SEVNLDAEFGXPKDD 20

RESULT 6
US-08-659-984A-18
: Sequence 18, Application US/08659984A
: Patent No. 5942400
: GENERAL INFORMATION:
: APPLICANT: Anderson, John P.
: APPLICANT: Sinha, Sukanto
: APPLICANT: Jacobson-Ciroak, Kirsten L.
: TITLE OF INVENTION: Assays for Detecting Beta-Secretase
: TITLE OF INVENTION: Inhibition
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Ctr., 8th Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/659,984A
: FILING DATE: 07-JUN-1996
: CLASSIFICATION: 436
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/485,152
: FILING DATE: 07-JUN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Heslin, James M.
: REGISTRATION NUMBER: 29,541
: REFERENCE/DOCKET NUMBER: 15270-0028100S
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-326-2400
: TELEFAX: 415-326-2422
: INFORMATION FOR SEQ ID NO: 18:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 21 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-08-659-984A-18

Query Match 22.8%; Score 33; DB 2; Length 21;
Best Local Similarity 50.0%; Pred. No. 39;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 14 SEISYGAEFNKD 25
|::: ||| |
Db 1 SEVNLDAEFRHD 12

```

RESULT 7  
US-08-802-981-113  
; Sequence 113, Application US/08802981  
; Patent No. 6037137  
; GENERAL INFORMATION:  
; APPLICANT: Komoriya, Akira  
; APPLICANT: Packard, Beverly S.  
; TITLE OF INVENTION: Compositions for the Detection of Enzyme  
; NUMBER OF SEQUENCES: 231  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/802,981  
; FILING DATE: 20-FEB-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hunter, Tom  
; REGISTRATION NUMBER: 38,498  
; REFERENCE/DOCKET NUMBER: 016865-000300US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 113:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; MOLECULE TYPE: linear  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 3  
; OTHER INFORMATION: /product= "Aib"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 4  
; OTHER INFORMATION: /product= "Acp"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 16  
; OTHER INFORMATION: /product= "Acp"  
US-08-802-981-113

Query Match  
Best Local Similarity 22.8%; Score 33; DB 3; Length 21;  
Matches 8; Conservative 53.3%; Pred. No. 39;  
Mismatches 2; Indels 3; Gaps 1;

OY 14 SEISVGAEFN--KDD 26  
||: ||| |||  
Db 6 SEVKDAEFGXPKDD 20

RESULT 8  
US-08-802-981-114  
; Sequence 114, Application US/08802981  
; Patent No. 6037137  
; GENERAL INFORMATION:  
; APPLICANT: Komoriya, Akira  
; APPLICANT: Packard, Beverly S.  
; TITLE OF INVENTION: Compositions for the Detection of Enzyme

; TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof  
; NUMBER OF SEQUENCES: 231  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/802,981  
; FILING DATE: 20-FEB-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hunter, Tom  
; REGISTRATION NUMBER: 38,498  
; REFERENCE/DOCKET NUMBER: 016865-000300US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 114:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; MOLECULE TYPE: linear  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 3  
; OTHER INFORMATION: /product= "Aib"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 4  
; OTHER INFORMATION: /product= "Acp"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 16  
; OTHER INFORMATION: /product= "Acp"  
US-08-802-981-114

Query Match  
Best Local Similarity 22.8%; Score 33; DB 3; Length 21;  
Matches 8; Conservative 53.3%; Pred. No. 39;  
Mismatches 2; Indels 3; Gaps 1;

OY 14 SEISVGAEFN--KDD 26  
||: ||| |||  
Db 6 SEVKDAEFGXPKDD 20

RESULT 9  
US-08-660-531-18  
; Sequence 18, Application US/08660531  
; Patent No. 6221645  
; GENERAL INFORMATION:  
; APPLICANT: Chrysler, Susanna M.S.  
; APPLICANT: Sinha, Sukanto  
; APPLICANT: Keim, Pamela S.  
; APPLICANT: Anderson, John P.  
; TITLE OF INVENTION: Beta-Secretase  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Ctr., 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA



ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/660,531  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/480,498  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Heslin, James M.  
REGISTRATION NUMBER: 29,541  
REFERENCE/DOCKET NUMBER: 15270-002210US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-660-531-18

Query Match 22.8%; Score 33; DB 4; Length 21;  
Best Local Similarity 50.0%; Pred. No. 39;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 14 SEISVGAEFNKD 25  
II:::III I  
DB 1 SEVNLDAEFFRHD 12

RESULT 10  
US-08-313-200-8  
Sequence 8, Application US/08313200  
Patent No. 5998153  
GENERAL INFORMATION:  
APPLICANT: Baker, James R.  
APPLICANT: Koenig, Ronald J.  
TITLE OF INVENTION: THYROID PEROXIDASE EPITOPIC REGIONS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FORSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/313,200  
FILING DATE: 08-NOV-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kanski, Antoinette F.  
REGISTRATION NUMBER: 34,202  
REFERENCE/DOCKET NUMBER: 20344-20658.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-313-200-8

Query Match 22.1%; Score 32; DB 2; Length 22;  
Best Local Similarity 54.5%; Pred. No. 61;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 14 SEISVGAEFNK 24  
I:::III II  
DB 10 SRDLTGALNKK 20

RESULT 11  
US-08-141-324-11  
Sequence 11, Application US/08141324  
Patent No. 5475097  
GENERAL INFORMATION:  
APPLICANT: Travis, James  
APPLICANT: Potempa, Jan S.  
APPLICANT: Bair, Philip J.  
APPLICANT: Pavloff, Nadine  
APPLICANT: Pike, Robert N.  
TITLE OF INVENTION: Lysine-specific Porphyromonas gingivalis  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee and Winer, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/141,324  
FILING DATE: 21-OCT-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 44-93  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 303-499-8080  
TELEFAX: 303-499-8089  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-141-324-11

Query Match 22.1%; Score 32; DB 1; Length 27;  
Best Local Similarity 50.0%; Pred. No. 79;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 14 SEISVGAEFNKD 25  
II:::III I  
DB 4 SEVVKDAEFFRHD 15

RESULT 12  
US-08-141-324-12  
; Sequence 12, Application US/08141324  
; Patent No. 5475097  
; GENERAL INFORMATION:  
; APPLICANT: Travis, James  
; APPLICANT: Potempa, Jan S.  
; APPLICANT: Barr, Philip J.  
; APPLICANT: Pavloff, Nadine  
; APPLICANT: Pike, Robert N.  
; TITLE OF INVENTION: Lysine-specific Porphyromonas gingivalis  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/141,324  
; FILING DATE: 21-OCT-1993  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ferber, Donna M.  
; REGISTRATION NUMBER: 33,878  
; REFERENCE/DOCKET NUMBER: 44-93  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303-499-8080  
; TELEFAX: 303-499-8089  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 27 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
US-08-141-324-12

Query Match 22.1% Score 32; DB 1; Length 27;  
Best Local Similarity 50.0%; Pred. No. 79;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 14 SEISVGAEPFND 25  
DB 4 SEVDLDAEPRHD 15

RESULT 13  
US-08-541-902-11  
; Sequence 11, Application US/08541902  
; Patent No. 5707620  
; GENERAL INFORMATION:  
; APPLICANT: Travis, James  
; APPLICANT: Potempa, Jan S.  
; APPLICANT: Barr, Philip J.  
; APPLICANT: Pavloff, Nadine  
; APPLICANT: Pike, Robert N.  
; TITLE OF INVENTION: Lysine-specific Porphyromonas gingivalis  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/541,902  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/141,324  
; FILING DATE: 21-OCT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ferber, Donna M.  
; REGISTRATION NUMBER: 33,878  
; REFERENCE/DOCKET NUMBER: 44-93  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303-499-8080  
; TELEFAX: 303-499-8089  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 27 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
US-08-541-902-11

Query Match 22.1% Score 32; DB 1; Length 27;  
Best Local Similarity 50.0%; Pred. No. 79;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 14 SEISVGAEPFND 25  
DB 4 SEVKMDAEPFHD 15

RESULT 14  
US-08-541-902-12  
; Sequence 12, Application US/08541902  
; Patent No. 5707620  
; GENERAL INFORMATION:  
; APPLICANT: Travis, James  
; APPLICANT: Potempa, Jan S.  
; APPLICANT: Barr, Philip J.  
; APPLICANT: Pavloff, Nadine  
; APPLICANT: Pike, Robert N.  
; TITLE OF INVENTION: Lysine-specific Porphyromonas gingivalis  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/541,902

FILING DATE: 435  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/141,324  
FILING DATE: 21-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Feider, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 44-93  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 303-499-8080  
TELEFAX: 303-499-8089  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-541-902-12

Query Match 22.1%; Score 32; DB 1; Length 27;  
Best Local Similarity 50.0%; Pred. No. 79;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 14 SEISVGAEFNKD 25  
||:|||||  
Db 4 SEVDLAEFRHD 15

RESULT 15  
US-08-324-301-3  
; Sequence 3, Application US/08324301  
; Patent No. 5597569  
; GENERAL INFORMATION:  
; APPLICANT: Siegall, Clay B.  
; APPLICANT: Siegall, Susan L.  
; APPLICANT: Marguardt, Hans  
; TITLE OF INVENTION: A NEW RIBOSOME-INACTIVATING PROTEIN  
; TITLE OF INVENTION: ISOLATED FROM THE PLANT BRYONICA DIOLOCA  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bristol-Myers Squibb Company  
; STREET: 3005 First Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98121  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/324,301  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/141,891  
; FILING DATE: 25-OCT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Poor, Brian W.  
; REGISTRATION NUMBER: 32,928  
; REFERENCE/DOCKET NUMBER: ON0109A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-728-4800  
; TELEFAX: 206-727-3601  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Bryonica dioica  
TISSUE TYPE: Root  
US-08-324-301-3

Query Match 21.4%; Score 31; DB 1; Length 19;  
Best Local Similarity 60.0%; Pred. No. 74;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 DVSLAESEIS 17  
::|||||  
Db 2 NIELGFEIS 11

RESULT 16  
US-08-706-741B-44  
; Sequence 44, Application US/08706741B  
; Patent No. 5955593  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, STANLEY J.  
; TITLE OF INVENTION: B3 INTERACTING DOMAIN DEATH AGONIST  
; NUMBER OF SEQUENCES: 88  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAVERKAMP, L.C.  
; STREET: 7733 FORSYTH BLVD., SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: USA  
; ZIP: 63146

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/706,741B  
FILING DATE: 09-SEP-1996  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 965017  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 727-5188  
TELEFAX: (314) 727-6092  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-706-741B-44

Query Match 20.7%; Score 30; DB 2; Length 13;  
Best Local Similarity 55.6%; Pred. No. 67;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CIGYHLDVS 10  
|:|:||||  
Db 5 CIGDEMDVS 13

RESULT 17  
US-08-924-695A-44  
; Sequence 44, Application US/08924695A  
; Patent No. 5998583

GENERAL INFORMATION:  
APPLICANT: KORSMEYER, STANLEY J.  
TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST  
NUMBER OF SEQUENCES: 88  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HOWELL & HAFFERKAMP, L.C.  
STREET: 7733 FORSYTH BLVD., SUITE 1400  
CITY: ST. LOUIS  
STATE: MISSOURI  
COUNTRY: USA  
ZIP: 63105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/924,695A  
FILING DATE: 09-SEP-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 971798  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 727-5188  
TELEFAX: (314) 727-6092  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-924-695A-44

Query Match 20.7%; Score 30; DB 2; Length 13;  
Best Local Similarity 55.6%; Pred. No. 67;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CIGYHLDVS 10  
|:|:|  
Db 5 CIGDEMVS 13

RESULT 18  
US-08-440-391-7  
Sequence 7, Application US/08440391  
Patent No. 5656725  
GENERAL INFORMATION:  
APPLICANT: CHITTENDEN, Thomas D.; and  
APPLICANT: LOTZ, Robert J.  
TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
MODULATE APOPTOSIS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,391  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: WIXON, HENRY N.

REGISTRATION NUMBER: 32,073  
REFERENCE/DOCKET NUMBER: 104322.147  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-942-8400  
TELEFAX: 202-942-8484  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-440-391-7

Query Match 20.7%; Score 30; DB 1; Length 15;  
Best Local Similarity 55.6%; Pred. No. 80;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CIGYHLDVS 10  
|:|:|  
Db 7 CIGDEMVS 15

RESULT 19  
US-08-440-391-32  
Sequence 32, Application US/08440391  
Patent No. 5656725  
GENERAL INFORMATION:  
APPLICANT: CHITTENDEN, Thomas D.; and  
APPLICANT: LOTZ, Robert J.  
TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
MODULATE APOPTOSIS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,391  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: WIXON, HENRY N.  
REGISTRATION NUMBER: 32,073  
REFERENCE/DOCKET NUMBER: 104322.147  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-942-8400  
TELEFAX: 202-942-8484  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-440-391-32

Query Match 20.7%; Score 30; DB 1; Length 15;  
Best Local Similarity 55.6%; Pred. No. 80;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CIGYHLDVS 10  
|:|:|  
Db 7 CIGDEMVS 15

RESULT 20  
US-08-248-839C-152  
; Sequence 152, Application US/08248839C  
; Patent No. 5843702  
; GENERAL INFORMATION:  
; APPLICANT: McConnell, David  
; APPLICANT: O'Kane, Kevin  
; TITLE OF INVENTION: A Gene Expression System  
; NUMBER OF SEQUENCES: 185  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 58437020 No. 5843702disk of No. 5843702th America, Inc.  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/248,839C  
; FILING DATE: 25-MAY-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gregg, Valeta A.  
; REGISTRATION NUMBER: 35,127  
; REFERENCE/DOCKET NUMBER: 3614.214-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 152:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Protein  
; US-08-248-839C-152

Query Match 20.7%; Score 30; DB 2; Length 15;  
Best Local Similarity 46.2%; Pred. No. 80;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 10 SLAFSEISVGAEF 22  
DB 3 SLAFPAVICSSEF 15

RESULT 21  
US-08-908-597A-7  
; Sequence 7, Application US/08908597A  
; Patent No. 5863795  
; GENERAL INFORMATION:  
; APPLICANT: CHITTENDEN, Thomas D.; and  
; APPLICANT: LUTZ, Robert J.  
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hale and Dorr  
; STREET: 1455 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/908,597A  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/440,391  
; FILING DATE: 12-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WIXON, HENRY N.  
; REGISTRATION NUMBER: 32,073  
; REFERENCE/DOCKET NUMBER: 104322.147  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-942-8400  
; TELEFAX: 202-942-8484  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-908-597A-7

Query Match 20.7%; Score 30; DB 2; Length 15;  
Best Local Similarity 55.6%; Pred. No. 80;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CIGYHLDVS 10  
DB 7 CIGDEMDS 15

RESULT 22  
US-08-908-597A-32  
; Sequence 32, Application US/08908597A  
; Patent No. 5863795  
; GENERAL INFORMATION:  
; APPLICANT: CHITTENDEN, Thomas D.; and  
; APPLICANT: LUTZ, Robert J.  
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hale and Dorr  
; STREET: 1455 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/908,597A  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/440,391  
; FILING DATE: 12-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WIXON, HENRY N.  
; REGISTRATION NUMBER: 32,073  
; REFERENCE/DOCKET NUMBER: 104322.147  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-942-8400  
; TELEFAX: 202-942-8484  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide

US-08-908-597A-32

## Query Match

Best Local Similarity 20.7%; Score 30; DB 2; Length 15;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;QY 2 CIGYHLDVS 10  
1:1:111  
DB 7 CIGDEMDS 15

## RESULT 23

US-09-236-385A-7

Sequence 7, Application US/09236385A

Patent No. 6221615

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

CORRESPONDENCE ADDRESSES:

ADDRESS: Hale and Dorr

STREET: 1455 Pennsylvania Avenue, N.W.

STATE: D.C.

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/236,385A

FILING DATE: 25-Jan-1999

CLASSIFICATION: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: WIXON, HENRY N.

REGISTRATION NUMBER: 32,073

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-942-8400

FAX: 202-942-8484

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-236-385A-7

Query Match

Best Local Similarity 20.7%; Score 30; DB 4; Length 15;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;QY 2 CIGYHLDVS 10  
1:1:111  
DB 7 CIGDEMDS 15

## RESULT 24

US-09-236-385A-32

Sequence 32, Application US/09236385A

Patent No. 6221615

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

MODULATE APOPTOSIS

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Hale and Dorr

STREET: 1455 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/236,385A

FILING DATE: 25-Jan-1999

CLASSIFICATION: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: WIXON, HENRY N.

REGISTRATION NUMBER: 32,073

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-942-8400

FAX: 202-942-8484

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 32:

US-09-236-385A-32

Query Match

Best Local Similarity 20.7%; Score 30; DB 4; Length 15;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;QY 2 CIGYHLDVS 10  
1:1:111  
DB 7 CIGDEMDS 15

## RESULT 25

PCT-US96-06122-7

Sequence 7, Application PC/TUS9606122

GENERAL INFORMATION:

APPLICANT: IMMUNOGEN, INC.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS

WHICH MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Hale and Dorr

STREET: 1455 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/06122

## CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/440,391

FILING DATE: 12-MAY-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: WIXON, HENRY N.

REGISTRATION NUMBER: 32,073

REFERENCE/DOCKET NUMBER: 104322.147PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-942-8400  
TELEFAX: 202-942-8484  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US96-06122-7

Query Match 20.7%; Score 30; DB 5; Length 15;  
Best Local Similarity 55.6%; Pred. No. 80;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 CIGYHLDVS 10  
I:I:III  
Db 7 CIGDEMDVS 15

Search completed: January 6, 2002, 09:13:06  
Job time: 424 sec

**THIS PAGE BLANK (USPTO)**